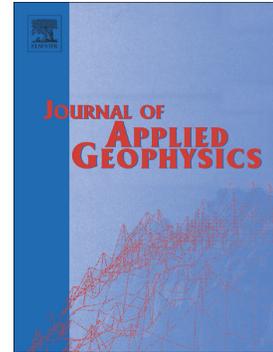


Journal Pre-proof

Joint interpretation of seismic refraction tomography and electrical resistivity tomography by cluster analysis to detect buried cavities

Carollo Alessandra, Capizzi Patrizia, Martorana Raffaele



PII: S0926-9851(20)30079-3

DOI: <https://doi.org/10.1016/j.jappgeo.2020.104069>

Reference: APPGEO 104069

To appear in: *Journal of Applied Geophysics*

Received date: 24 January 2020

Revised date: 12 May 2020

Accepted date: 13 May 2020

Please cite this article as: C. Alessandra, C. Patrizia and M. Raffaele, Joint interpretation of seismic refraction tomography and electrical resistivity tomography by cluster analysis to detect buried cavities, *Journal of Applied Geophysics* (2019), <https://doi.org/10.1016/j.jappgeo.2020.104069>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.

Joint interpretation of seismic refraction tomography and electrical resistivity tomography by cluster analysis to detect buried cavities

Carollo Alessandra¹, Capizzi Patrizia¹, Martorana Raffaele¹

¹ Università degli Studi di Palermo, Dipartimento delle Scienze della Terra e del Mare (DiSTeM), Palermo, Italy

Corresponding Author: alessandra.carollo02@unipa.it

ABSTRACT

In the last few years, the geophysical methods of seismic refraction tomography (SRT) and electrical resistivity tomography (ERT) are among the most used geophysical techniques for the reconstruction of subsoil geometries, for the investigation of underground cavities and also for the archaeological prospecting. However, the main disadvantage of each geophysical method is the difficulty of final interpretation of the data. In order to eliminate artifacts and generally improve the reliability and accuracy of geophysical interpretation, it is useful to perform a joint approach of different geophysical methods, also introducing a priori information. In this work, it is shown the integrated study of seismic refraction tomography and electrical resistivity tomography techniques, the two geophysical methods are tested on both synthetic and real data and the integration of data is useful in detecting buried cavities and also evaluate their geometric characteristics. Likelihood parameters has been defined and tested, in order to help recognizing voids from other lithological structures. Finally, a statistical approach based on cluster analysis of the P-wave velocity, the density of the seismic rays and the electrical resistivity of the synthetic and experimental models was used. Multi-space cluster distribution maps were built, allowing to better define and interpret the anomalies of the subsoil.

Keywords: SRT; ERT; Joint interpretation; K-means cluster analysis; Modelling; Cavity.

1. Introduction

The presence of cavities buried (voids like tunnels, underground canals, underground mines, caves, tombs) in highly populated areas creates serious problems in terms of hazard and risk for the

stability of the infrastructure and individual protection. It is therefore essential to identify these areas to minimize the geological risk and implement an optimal plan for the use of the soil. Several case studies, from around the world, show the effectiveness of geophysical methods in obtaining information on the structures and geometries present in the subsoil and constitute an excellent tool for the detection of buried cavities (Imposa et al., 2018; Maraió et al., 2015; Orfanos et al., 2008). The determination of dimension and depth of the cavities and of the contrast in physical properties between the cavity and surrounding media are fundamental aspects for understanding the validity of a particular geophysical method (Fasani et al., 2013). Until today, different geophysical prospecting methods have been used for the detection of underground voids, and their success is strongly influenced by their non-invasiveness, the ability to pre-investigate and data resolution. The ambiguities of the inverse models suggest the integrated use of different geophysical methodologies in order to obtain information of different nature on independent parameters and to correlate them during the interpretative process.

It is also essential that the geophysical survey is designed and chosen on the basis of a priori information of the geological context, making it possible to obtain a rough estimate of depth and size of the voids to evaluate any differences in the properties physico-mechanical properties of the materials involved (Cardarelli et al., 2010). Today electrical resistivity tomography (ERT) is the most widely method used for the detection of underground voids (Capizzi et al., 2005) and also for archaeological research (Cardarelli et al., 2008).

However, this technique is less sensitive to electrical and electromagnetic disturbances when compared to other geophysical methods (Van Schoor, 2002; McDonald et al., 2003; Cardarelli et al., 2008). Consequently, in recent years, the technique of electrical tomography has been joined more and more often to the technique of seismic refraction tomography (SRT) in order to obtain more robust interpretations (Sheehan et al., 2005; Riddle et al., 2010). The use of combining electrical and seismic data derives from the observation that there is an indirect correlation between the electrical resistivity and the velocity of the seismic waves, through a direct correlation of both parameters with the porosity (Archie, 1942; Wyllie et al., 1956), caused by the influence on both these geophysical parameters by the structure of the pores of the materials (Meju et al., 2003). This is the main assumption for a joint inversion of seismic refraction and electrical resistivity tomography. This can be very useful for the identification of cavities and tunnels (Dahlin et al., 1999), because it allows to reduce the intrinsic ambiguities of each method (Linder et al., 2010), at the same time improving the resolution and compatibility of the velocity and resistivity models (Gallardo and Meju, 2004).

A correct management of a multiparametric dataset must take into account the difficulty of defining valid theoretical or at least empirical relations between the parameters. These are in fact influenced by the great variability of the chemical and physical conditions within the Earth (Lees and Van Decar, 1991). It is therefore necessary to allow a certain degree of independence between each type of observed data, modeling separately each geophysical parameter in the forward problem. In this way it is possible to impose mathematical constraints that guide the joint inversion towards models with similar spatial distributions of anomalies (Gallardo and Meju, 2003, 2004, 2007, 2011). However, often, the parameters considered (for example, seismic velocity and electrical resistivity) have different distributions in the subsoil and, in these cases, the conjugate inversions can lead to serious misinterpretations.

An approach that is alternative to the joint inversion involves the use of post-inversion techniques for univariate independent models, in order to find relationships between the different observable parameters to identify different lithological structures. Following this approach, local empirical relationships between different parameters can be defined.

Cluster analysis is a multivariate analysis technique in which statistical units can be combined using an optimization criterion, minimizing the parametric distance within each group and at the same time maximizing it. The parametric distance is quantified by measures of similarity and / or dissimilarity between defined statistical units. Thus, with the cluster analysis method, it is possible to identify within a set of objects some subsets called clusters, which tend to be similar within them, based on their level of similarity. In this way, each cluster contains elements that are homogeneous to each other and which have a high internal coherence (ie, minimum intra-cluster distance) and high external heterogeneity (ie, maximum intercluster distance) (Barbarito, 1999). In general, clustering is an excellent tool whenever you want to catalog a certain amount of information in meaningful and manageable groups. Over the years, clustering methods have been applied to a wide variety of research fields, the scholar Hartigan (1975) provides a comprehensive list of the numerous published studies reporting the results of various cluster analysis applications.

Several authors used the cluster analysis between the seismic velocity and the electrical resistivity distribution (i.e. Dell'Aversana, 2001; Gallardo and Meju, 2003; Bottari et al., 2018a). K-means cluster analysis has often been applied in order to correlate different physical properties between them, for example seismic velocity and electrical resistivity (Bottari et al., 2018b), velocity and attenuation of the electromagnetic waves in georadar data (Tronicke et al., 2004), georadar attenuation and magnetic field intensity (Scudero et al., 2018), datasets of seismic refraction tomography (SRT) and controlled-source audio magnetotelluric (CSAMT) (Di Giuseppe et al., 2014).

The use of post-inversion techniques applied to ERT and SRT data for the investigation of cavities has been successfully proposed using different types of clustering approaches, in order to facilitate the phase of interpretation by providing a clearer imaging of the representation of the subsoil (Meju et al., 2003; Gallardo and Meju, 2004; Orfanos and Apostolopoulos, 2013; Di Giuseppe et al., 2014; Kotyrba and Schmidt, 2014; Hellman et al., 2017).

In this paper the application of an algorithm based on non-hierarchical cluster analysis (k-means) is proposed, together with the use of likelihood parameters, for the joint interpretation of SRT and ERT data, acquired with identical sensor array (electrodes and geophones). The methodology is tested on synthetic models that represent cavities at different abscissas and depths, together with structures of similar shape but with physical parameters not indicative of the presence of cavities, and the advantages are discussed. Finally, the proposed approach is tested on a site of limestone quarries, on which a speleological survey was previously made.

2. Materials and methods

2.1 Choice of the synthetic models

Synthetic modeling means creating a simplified representative model of a subsoil portion. The most effective way to represent with a good approximation the complex distributions of a physical parameter in the subsoil is to perform a discretization in homogeneous cells, each characterized by constant values of the physical parameters considered. In this way the synthetic model is not a complete representation of the real model, since it is characterized and described by a finite number of elements, but it is possible to decide the resolution degree of the model by varying the dimensions of the cells. By solving a forward problem, it is possible to calculate the predicted data according to the analyzed physical laws (elastic wave propagation, electric field, etc.). The forward modeling allows to study potentialities and limits of the analyzed methodology. Similarly with other authors (Cardarelli et al., 2010; Fasani et al., 2013), in this study synthetic models that represented different cavities were considered, parametrizing the velocity of the pressure seismic waves and the electrical resistivity. These synthetic models were used to test the properties, efficacy and limits of the joint use of seismic refraction tomography (SRT) and electrical resistivity tomography (ERT) for the detection of sub-surface cavities, using post-inversion cluster analysis techniques. Moreover, a detailed synthetic modeling was also performed, based on a real situation in which high-resolution speleological surveys were available, in order to compare predicted and experimental results, to optimize the inversion and interpretation phase. All the synthetic models of seismic refraction tomography were created by Surfer® software (Golden Software), while the synthetic data were calculated, processed and inverted with the Rayfract® software (Intelligent

Resource, Inc.). The resistivity models were created with the RES2DMOD® software and inverted with the RES2DINV® software.

2D synthetic models, synthetic models were created with different number of cavity and blocks of highly cohesive lithological material (high seismic velocity and resistivity values). The electrical resistivity models were constructed using a value of 10000 Ωm for cavities and 500 Ωm for the background material. While, refraction seismic models were constructed (Rayfract 2010a; Rayfract 2010b) using a value of 330 m/s for cavities and 5000 m/s to simulate the presence of a highly cohesive material (box size 6 m x 6 m).. The background of the seismic models has been divided into layers to create a vertical velocity gradient, with v_p values between 400 m/s near the surface up to 3000 m/s in depth. The layers have variable thicknesses: for the first four layers near surface the thickness is 2 m, the underlying layer instead has a thickness of 7 m (bottom at depth of 15 m) and the last layer reaches the depth of 30 m coinciding with lower boundary of the model. The difference between seismic and electrical background models derives from the observation that stone rocks often show surface layers of alterations that cause a decrease in seismic velocity that does not correspond to a measurable variation in their high resistivity values.

The synthetic seismic coverage and refraction tomographies were created, processed and inverted with Rayfract® software (Intelligent Resource, Inc.), while the synthetic resistivity models have been created and inverted by the RES2DMOD® and the RES2DINV® software, respectively.

The synthetic models (Fig. 1) were created starting from the same background subsoil and inserting in it anomalous zones that represented empty or filled cavities or different anomalous zones.

Model A (Fig. 1 a), represents a single square section cavity. In the middle of the section a square-shaped has been inserted (6 m x 6 m), having its top at depth of 6 m. This anomaly simulates an air-filled cavity characterized by a P-waves velocity of 330 m/s and a resistivity of $10^5 \Omega\text{m}$.

Model B (Fig. 1 b) represents three cavities and shows three anomalous square-shaped zones (6 m x 6 m) located between 29 m and 35 m, between 45 m and 51 m, and between 61 m and 67 m respect to left limit of the profile, and having the top at depths respectively of 4 m, 6 m and 8 m. The three anomalies are characterized by P-wave velocity equal to 330 m/s and electrical resistivity equal to $10^5 \Omega\text{m}$.

Model C represents two cavities and a compact calcarenitic block (Fig. 1 c). It presents the same three anomalous square-shaped zones as the model B. However, the most superficial block (on the left) represents a compact calcarenitic block with a P-wave velocity of 5000 m/s and a resistivity of 5000 Ωm . The other two deeper blocks, on the other hand, represent two cavities with P-wave velocities equal to 330 m/s and resistivity equal to $10^5 \Omega\text{m}$.

Model D, similarly to the model C, represents two cavities and a compact calcarenitic block (Fig. 1 d). The difference is that the compact calcarenitic block, with $v_p = 5000$ m/s and $\rho = 5000$ Ωm , is the square-shaped zone in the middle of the model, and the left and right square-shaped zone represent air-filled cavities ($v_p = 330$ m/s and $\rho = 10^5$ Ωm).

Model E, similarly to models C and D, represents two cavities and a compact calcarenitic block (Fig. 1 e). In this case the calcarenitic block, with $v_p = 5000$ m/s and $\rho = 5000$ Ωm , is the square-shaped zone in the right of the model, and the left and middle square-shaped zone represent air-filled cavities ($v_p = 330$ m/s and $\rho = 10^5$ Ωm).

Finally, model F (Fig. 1 f), is a 2D section obtained by a detailed synthetic modeling based on high-resolution speleological survey carried out the Foderà Quarry (Fig. 2 a) near Marsala, in Sicily (Fig. 2 b), that has been chosen as field test site in order to compare predicted and experimental results. A map of the Quarry (Fig. 2 c) was plotted thanks to a 3D speleological survey carried out in 2016. Starting from this survey a section of the subsoil (A-B) was obtained and used as a base to simulate the synthetic model F; the profile corresponds to the location of the electric and seismic tomography line carried out in field.

Model F has been set on the same background considered for the other models. Moreover, it must be said that the range of seismic waves velocity and resistivity values has been chosen according to the characteristic lithology of the quarry area: being a compact calcarenitic rock, v_p reaches a maximum value of 3000 m/s and ρ is about 500 Ωm .

The section obtained from the speleological survey intersects four cavities that are sketched in the synthetic model, all at 6 m depth, with variable thicknesses that reach a maximum depth of 21 m from the countryside level. Starting from the left, the first cavity is located from 28 m to 34 m from the origin of the profile, it has a thickness of 13 m and reaches a depth of 19 m; the second cavity is located from 45 m to 50 m, it has a thickness of 11 m and reaches a depth of 17 m; the third cavity, located from 58 m to 65 m, has an irregular shape and thickness ranging from a minimum of 2 m to a maximum of 8 m, it reaches a depth of 10 m; the fourth cavity located from 70 m to 100 m, also has irregular shape and variable thickness up to a maximum of 15 m, reaching a depth of 21 m. The values of the geophysical parameters assigned to these cavities are the same as for the other synthetic models: $v_p = 330$ m/s and $\rho = 10^5$ Ωm .

At the beginning, the model obtained from the speleological survey did not include the fourth cavity on the right, added later on the basis of the results of the field surveys and subsequently identified by speleologists. This is a clear example of how in field the choice of position and total length of the array does not always turn out to be optimal.

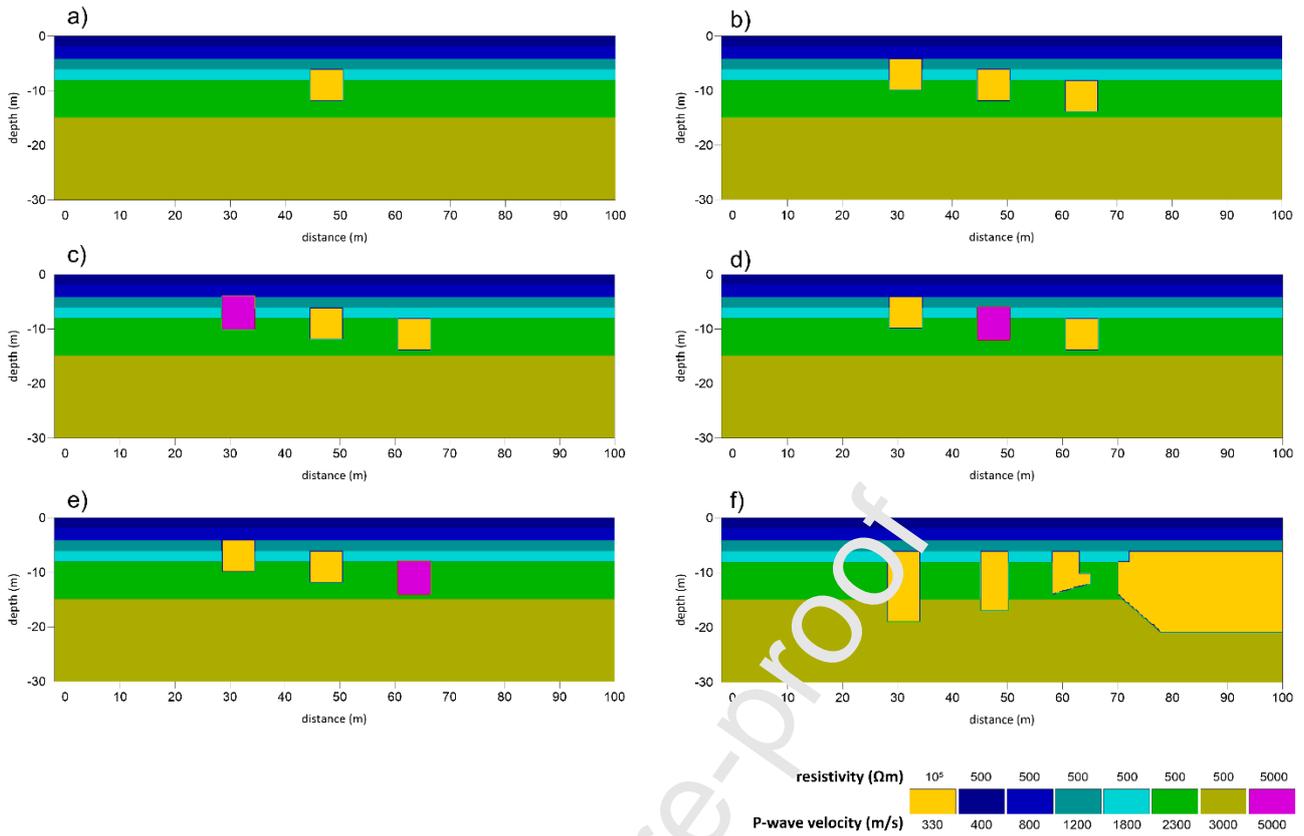


Fig. 1. Simulated models representing different combinations of cavities ($v_p = 330$ m/s, $\rho = 10^5$ Ωm) and compact calcarenitic blocks ($v_p = 5000$ m/s, $\rho = 5000$ Ωm): a) model A representing a buried cavity with its top at depth of 6 m; b) model B representing three cavities, with their top at depths respectively of 4 m, 6 m and 8 m; c) model C representing a calcarenitic compact block (left) with top at a depth of 4 m, and two cavities (middle and right), with their top at depths respectively of 6 m and 8 m; d) model D representing a calcarenitic compact block (middle) with top at a depth of 6 m, and two cavities (left and right) with their top at depths respectively of 6 m and 8 m; e) model E representing a calcarenitic compact block (right) with top at a depth of 8 m, and two cavities (left and middle) with their top at depths respectively of 4 m and 6 m; f) model F representing the Foderà Quarry section, including four cavities of irregular shapes, with top at the same depth of 6 m.

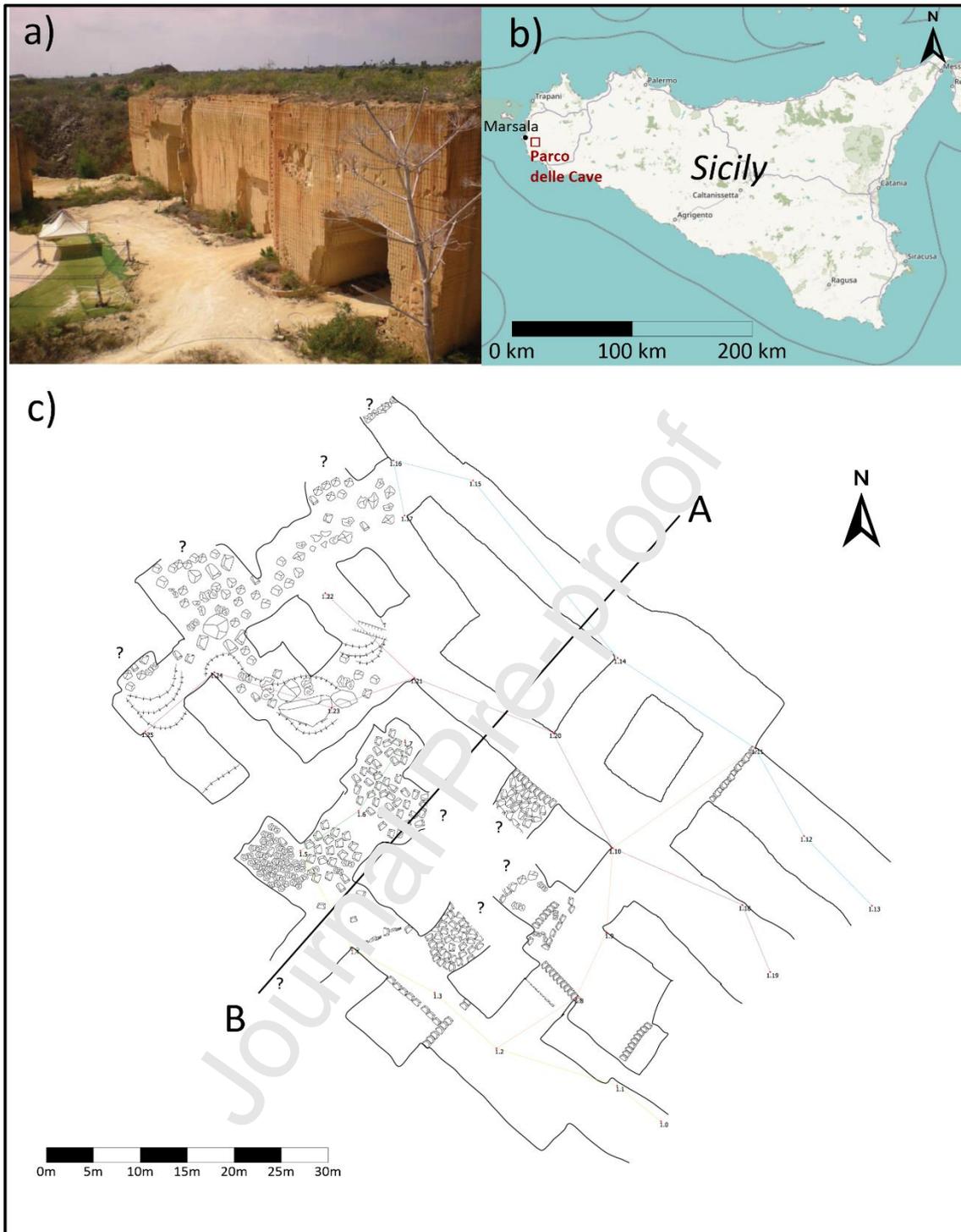


Fig. 2. a) A picture of the Foderà Quarry, in which three tunnel entrances are visible; b) Ubication of “Parco delle Cave” near the city of Marsala, in Sicily; c) Map of the Foderà Quarry, obtained after a 3D speleological survey, showing the trace A-B of the joint SRT and ERT survey.

2.2 Simulating joint-interpreted seismic refraction and electrical resistivity tomographies

Starting from the aforementioned models, the execution of joint-interpreted SRT and ERT were simulated, considering the same positions for the geophones and the electrodes, coinciding with the surface nodes of the model cells.

For the seismic refraction survey 48 geophones were considered, 2 meters spaced, and 25 shots distributed along equal distances of 2 m, with left and right offset equal to 2 m, for a total length of the seismic line equal to 100 m. This array should allow a good resolution of the inverse model, with a quite homogeneous coverage.

Seismic refraction data was inverted using Rayfract® (Intelligent Resource, Inc.) ver. 3.12. It uses the WET (Wavepath Eikonal Traveltime) method (Schuster and Quintus-Bosz, 1993; Rohdewald, 2016). The same set of damping and inversion parameters and the same cell sizes were used for all the data sets related to different synthetic models, aiming to obtain a reliable comparison among all the tomographies. A manual picking of each track was made to estimate the first arrival times of P-waves. As a first step, synthetic traveltimes data has been inverted using the Delta-t-V method (Gibson et al., 1979) which automatically generates an initial 1D model directly from the chosen traveltimes (Gebrande and Miller, 1985; Rohdewald, 2011), to identify small features and velocity inversions. Subsequently, an iterative back-projection of the wavefronts using a finite-difference solution to the eikonal equation (Qin et al., 1992) was applied to the Delta-t-V results.

The electrical resistivity survey was simulated considering 48 electrodes, 2 meters spaced, for a total length of 94 m. The inline dipole-dipole array was considered, with dipole length a ranging from 1 to 4 times the electrode spacing and dipole-dipole distance n ranging from 1 to 8 times the electrode spacing. This data set comprises a total of 945 measures and is ideal for making the most of the capabilities of modern multi-channel resistivity-meters. In fact, it allows a good resolution and reliability of the survey, according with a not too long acquiring time due to an adequately high ratio between the number of measures and that of current dipoles (Martorana et al., 2017a).

Inversion of predicted apparent resistivity data was performed using the RES2DINV™ and RES3DINV™ software applications (Loke, 2013). The same set of damping and inversion parameters and the same cell sizes were used for all the ERTs, in order to correctly compare the results. Considering the large contrast in the apparent resistivity relative to adjacent measurements, a further model refinement was performed, with mesh size equal to a half of the electrode distance. Moreover, a mesh grid with two nodes per electrode spacing was chosen in the forward modelling routine, in order to accurately calculate the theoretical values. An L_1 norm, iteratively reweighted least squares method (Wolke and Schwetlick, 1988) was used to obtain models capable to highlight sharp boundaries with a high resistivity contrast. We consider this to be the best choice if we want to correctly outline the shape of a cavity.

2.3 Likelihood parameters for cavity presence

The size, the depth of burial of the cavity and the contrast of the physical properties between the host rock and the cavity, are aspects that play an important role for the evaluation of the geophysical method for the detection of cavities. ERT is a well-established and widely used method to detect cavities; however, the information obtained from SRT is capable of eliminating some potential ambiguities in the interpretation of ERT data. Both geophysical methods have been applied to both synthetic and real data, allowing to identify and evaluate the geometric characteristics of the cavities analyzed (Cardarelli et al., 2010). In general, if the combined inversion of ERT and SRT data shows a region with high resistivity and low seismic velocity, this is considered as a region in which a vacuum is present (Sheehan et al., 2005). Vice versa, if an anomaly shows low resistivity and high seismic velocity, the presence of the vacuum in the subsoil can be excluded. In fact, therefore, the two correlated physical properties whose high values can indicate the presence of cavities are the electrical resistivity and the seismic slowness s_p , inverse of the seismic velocity v_p .

To detect a body in the subsoil it is also necessary to have adequate coverage of seismic rays. In the case of a cavity the seismic rays will concentrate in the area around the cavity and will decrease at the cavity itself. Moreover, in order for the SRT to be able to detect a cavity, it is essential that the seismic rays penetrate a depth greater than that of the cavity base, to be thus critically refracted to the surface, if the velocities of the underlying layers increase with depth.

With the above considerations in mind, we choose two likelihood parameters that could highlight the presence of cavities. These parameters have been estimated starting from the normalization of the physical quantities considered. The normalized values of each quantity is calculated as follows:

$$N(x) = (x - x_{\min}) / (x_{\max} - x_{\min}), \quad (1)$$

where x is the physical quantity involved and x_{\min} and x_{\max} are the minimum and maximum values in the inverse model.

The first parameter, here named P_1 , is the product between the normalized logarithm of resistivity ρ and the normalized P-wave slowness $s_p = 1 / v_p$:

$$P_1(\rho, s_p) = N(\text{Log } \rho) * N(s_p). \quad (2)$$

The second one, here named P_2 , is the ratio between P_1 and the normalized seismic ray density d :

$$P_2(\rho, s_p, d) = N(\text{Log } \rho) * N(s_p) / N(d). \quad (3)$$

These likelihood parameters are able to help to differentiate with more precision the cavity (values close to 1) from other structures that could imply high values of electrical resistivity or high values or seismic slowness but not both these contemporary events. P_1 and P_2 were calculated for each cell of the inverse models.

2.4 Cluster analysis

In cluster analysis there are two types of classification algorithms: hierarchical algorithms and non-hierarchical algorithms. As for hierarchical algorithms, each statistical unit is part of a larger group, which is itself contained in an even larger group, until a single group containing the entire population is obtained. For non-hierarchical algorithms it is necessary to decide the number of clusters to be obtained and, based on this number, the algorithm tries to obtain the best possible grouping. Generally, cluster analysis does not require an a priori model of interpretation (Fabbris, 1983). The centroid-based clustering technique can be theoretically considered as an optimization problem that proceeds according to an iterative process: in the preliminary phase it is necessary to find the k cluster centers and to associate each object to its cluster so that the distances squared by the centroid of the cluster are kept to a minimum; subsequently it is needed to calculate the new averages to be the centroids of the observations in the new clusters. The algorithm reaches its final optimum (local) when the assignments no longer change and the clustering procedure is interrupted; however, in some cases there is no certainty of finding the global optimum. Typically, centroid-based algorithms require to identify and specify in advance the number of clusters, k , and the coordinates of the initial centroid. This aspect is considered one of the major problems of these algorithms, because an inappropriate choice of k can produce poor results. It is also important to do diagnostic research to determine the number of clusters in the data set.

In this work, a statistical approach based on non-hierarchical cluster analysis (k -means) has been applied, as it is the most efficient approach in computational terms, less influenced by the anomalous values and allows a statistical unit to change your cluster during the iterative process (Martorana et al., 2017). Moreover, the final results obtained by this non-hierarchical approach are certainly simpler to interpret because they are represented by a single partition, which is interpreted on the basis of its hypotheses. The idea of using cluster analysis in geophysical exploration is that the structures in the subsoil are easily distinguishable using a set of different physical properties compared to a single property. To perform cluster analysis on data, an algorithm implemented in MATLAB® is used. This algorithm aims to group units into classes in a single solution, based on optimization criteria, minimizing variance within the cluster and at the same time maximizing the distance between clusters. The data acquired in this study are favorable to cluster analysis, since the values of the propagation velocity of the wave P , the density of the seismic rays and the electrical resistivity are available on coincident sections with comparable resolution. The distance of each element from the initial nucleus and from the nuclei obtained after each iteration was calculated from the weighted sum of the Euclidean distances of all the parameters considered:

$$D = a \sqrt{(dx^2 + dy^2)} + b \sqrt{(dC^2)} + c \sqrt{(dV^2)} + d \sqrt{(d\rho^2)} \quad (4)$$

where a , b , c and d are weights, dx , dy , dC , dV and $d\rho$ are respectively the differences between the spatial coordinates x and y , the density of the seismic rays, the propagation velocity of the P wave and the electrical resistivity. The choice of the optimal number of clusters (k) have been optimized maximizing R^2 parameter, taking into account the intra-cluster (DEVin) and inter-cluster (DEVout) variances:

$$R^2 = \text{DEVout}/\text{DEVt}, \quad \text{DEVt} = \text{DEVin} + \text{DEVout} = 1 \quad (5)$$

Passing from $k+1$ to k groups (i.e. the aggregation phase) DEVin increases, while obviously DEVout decreases. At each iteration groups are aggregated among them so there is the least increase in variance within the groups.

The first phase of elaboration consists in the independent derivation of the resistivity inversion model and the inversion model of seismic velocities; the values obtained are processed by the algorithm proceeding with the first iteration by choosing the coordinates of these quantities, to be represented in a common graph x - z , where each node of this grid has a value of electrical resistivity, seismic velocity and seismic ray density. At the end of the iterations, an x - z profile is generated where the nodes belonging to the profile are colored according to the starting cluster.

With clustering it is also possible to analyze the different parameters by plotting them within a same 3D multiparametric graph, delimiting the limits of the data group areas and facilitating the reading of the detected information and obtaining a more complete view. Generally, the graphical representation of the cluster occurs through the use of a color code that encloses a group of clusters with similar values.

3. Synthetic test results

3.1 Model A

In Fig. 3 the results of the inversion of simulated data obtained starting from the model A are showed. The inversion processing of the seismic data required 20 iterations to reach a misfit value of $4 \mu\text{s}$ (Schuster, 1993). The distribution of the ray path density (Fig. 3 a) shows a decrease at the cavity and an increase at the lower limit of the cavity. The Seismic Refraction Tomography (Fig. 3 b), shows an evident low velocity at the cavity, even if the inverted values are greater than those of the theoretical model. The electrical resistivity tomography (Fig. 3 c) show high values of resistivity at the cavity, low values even if lower than those of the theoretical model. Both SRT and ERT identify the cavity, even if with a wider anomaly in ERT (Riddle et al., 2010).

The distribution of the likelihood parameter P_1 (Fig. 3 d) and P_2 (Fig. 3 e) shows that the cavity area is well delimited, with both values close to one at the cavity.

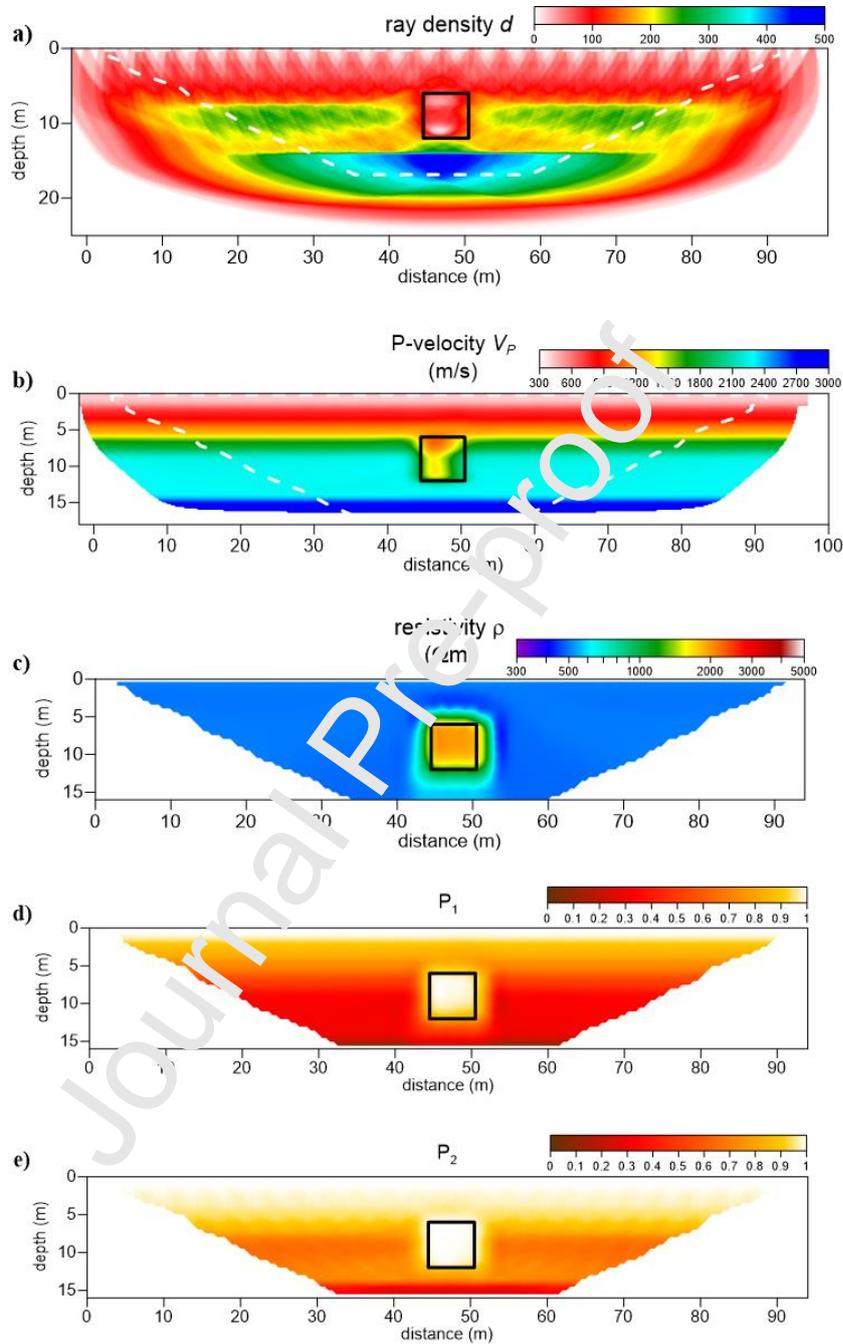


Fig. 3. Model A: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities are indicated by black solid lines. The joint interpreted section is contoured by a dashed white line.

The cluster analysis performed considering the geophysical parameters of the electrical resistivity, the propagation velocity of the P waves and the coverage of the seismic beams, assigns each cell of the model to a specific cluster characterized by similar trends of the aforementioned

parameters. Depending on the choice of number k of the classes, the algorithm returned different patterns (Fig. 4). The cavity is well identified if $k \geq 3$. If $k = 5$ we can also differentiate the vertical gradient of seismic velocity.

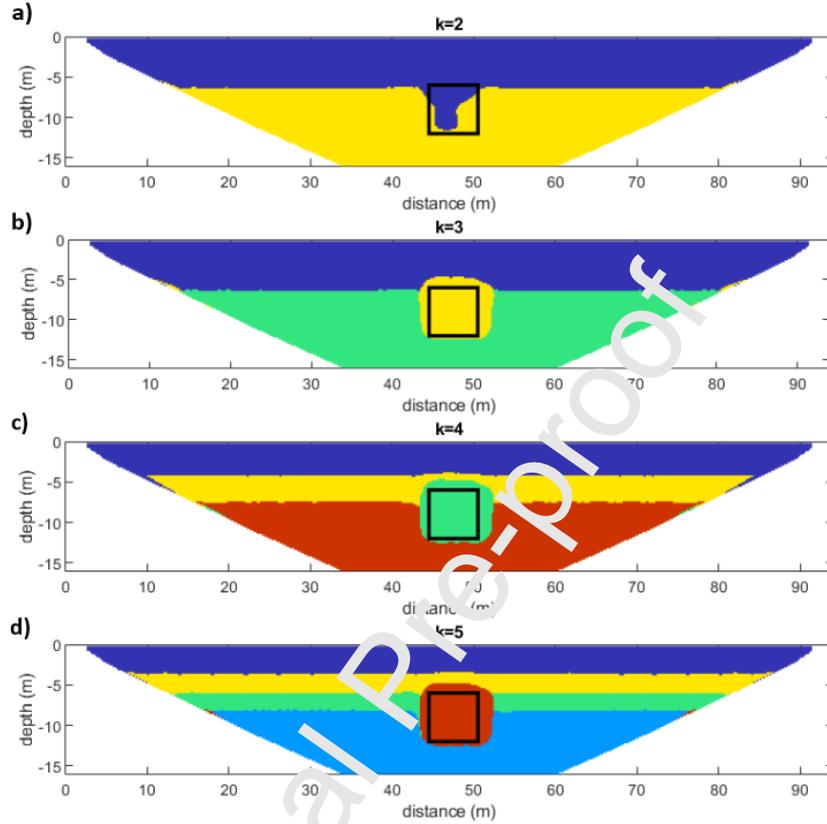


Fig. 4. Model A: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

3.2 Model B

The results of the inversion of simulated data obtained starting from the model B are showed in Fig. 5. After 20 iterations a misfit of $6 \mu\text{s}$ was obtained between predicted and measured travel times of P-waves. The distribution of the ray path density (Fig. 5 a) shows evident decreases corresponding to the three cavity zones located at the different depths; while a high ray path density zone, between 14 m and 20 m of depth, is placed underneath the cavities marking the lower limits of these. The SRT (Fig. 5 b) shows a decrease in velocity at the three cavities. ERT (Fig. 5 c) shows high resistivity values that however decrease as the depth of the cavities increases; the deeper cavity, located at 8 m depth, could be identified with more difficulty because of a not so high resistivity contrast and more smoothed boundaries; low resistivity values delineate the volume surrounding the cavities.

The distribution of the likelihood parameter P_1 (Fig. 5 d) and P_2 (Fig. 5 e) are able to discriminate the three cavities to which low values of conductivity, P-wave velocity and ray path density correspond. In this case the identification of the deeper cavity is more accurate than the inversion of the only SRT or ERT, especially if we look at the distribution of P_2 .

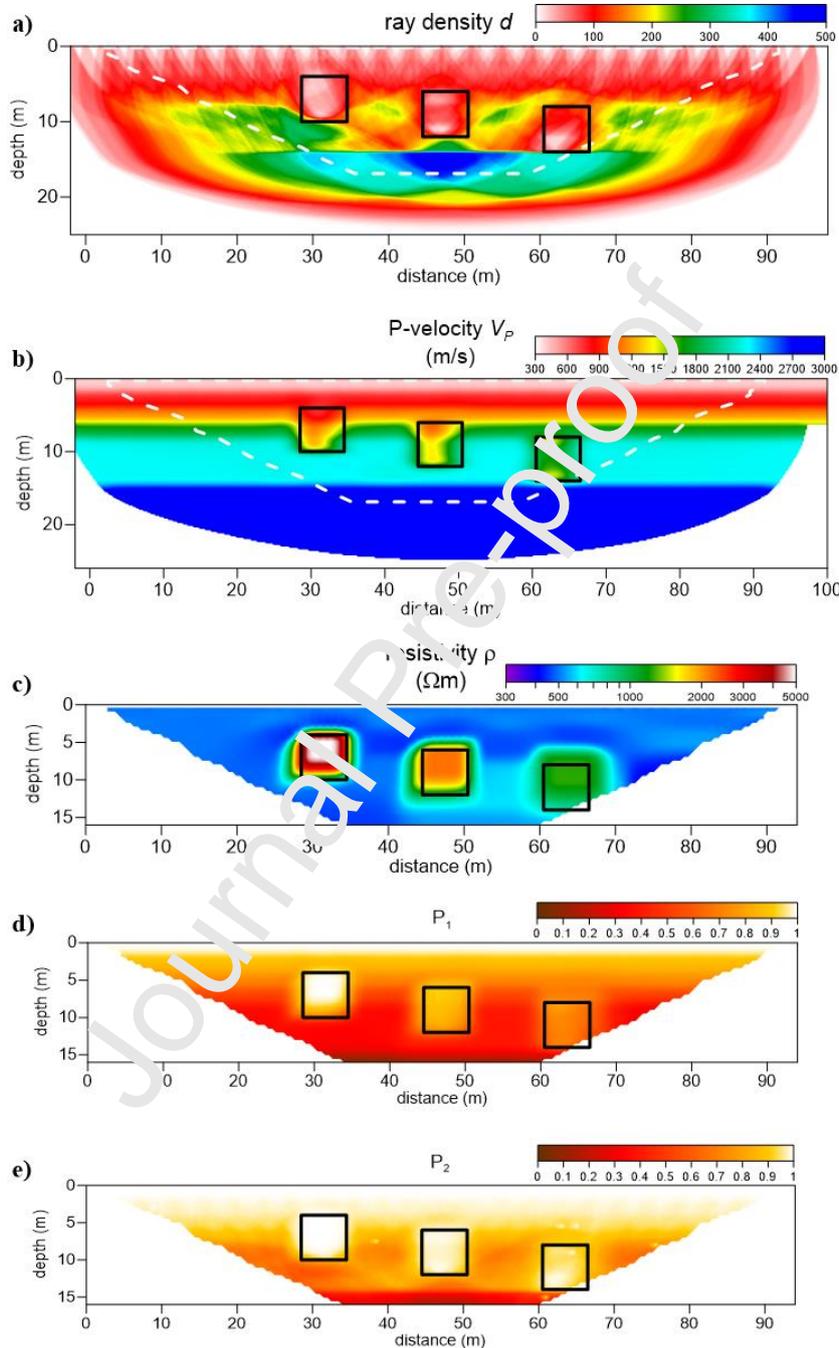


Fig. 5. Model B: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities are indicated by black solid lines. The joint interpreted section is contoured by a dashed white line.

The cluster analysis is obviously dependent on the choice of the value of k classes. With $k = 2$ (Fig. 6 a) the image discriminates two large clusters the delimit quite well only the bottom of the

cavities. For $k = 3$ (Fig. 6 b) there is a cluster that delimit the area underneath the cavities, yet information about the cavities is insufficient or even missing for the deeper cavity.

If $k = 4$ (Fig. 6 c) all the cavities are well isolated, identified with a single cluster. With $k = 5$ (Fig. 6 d) another cluster defines the outer edges of the image.

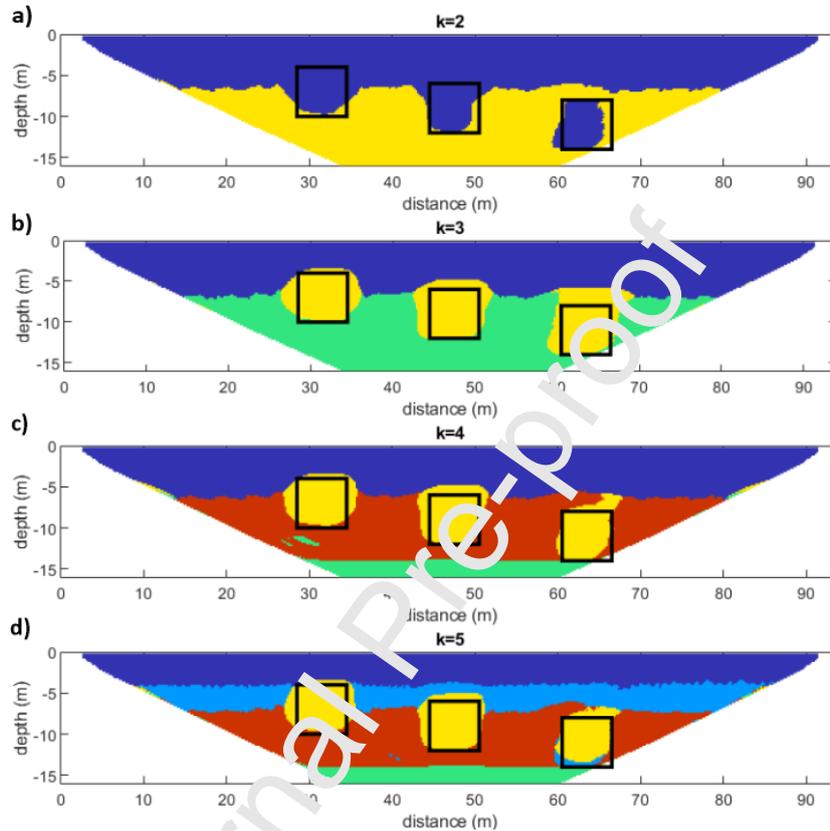


Fig. 6. Model B: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

3.3 Model C

The results of the inversion of simulated data obtained starting from the model C are showed in Fig. 7. After 50 iterations a misfit of $12 \mu s$ was obtained between predicted and measured travel times of P-waves. There is a decrease in the ray path density (Fig. 7 a), corresponding to the two cavities; while a high density area is evident at 4 m of depth (corresponding to the simulated compact calcarenitic block), but it also propagates at depths between 11 m and 20 m below the cavities, delineating the lower limits. SRT (Fig. 7 b) shows the decrease of v_p at 6 m and at 8 m of depth where the cavities are located; a high seismic velocity zone is showed in depth and tends to expand towards the surface in correspondence to the simulated compact calcarenitic block is located. ERT (Fig. 7 c) shows high resistivity values for the cavity at a depth of 6 m and low values and strongly blunted limits for the cavity located at 8 m depth. It also well highlights the

calcarenic compact block at 4 m depth, with higher resistivity values compared to adjacent deeper cavities. ERT therefore fails to differentiate the calcarenitic compact block from the cavities.

In the anomalous zones the likelihood parameter P_1 (Fig. 7 d) shows values between 0.7 and 1 progressively decreasing with depth and with increasingly blunt limits; in the same zones the range of variation of P_2 (Fig. 7 e) is smaller (0.85-1). In this case the parameters P_1 and P_2 do not allow to discriminate between the cavities and the compact calcarenitic block.

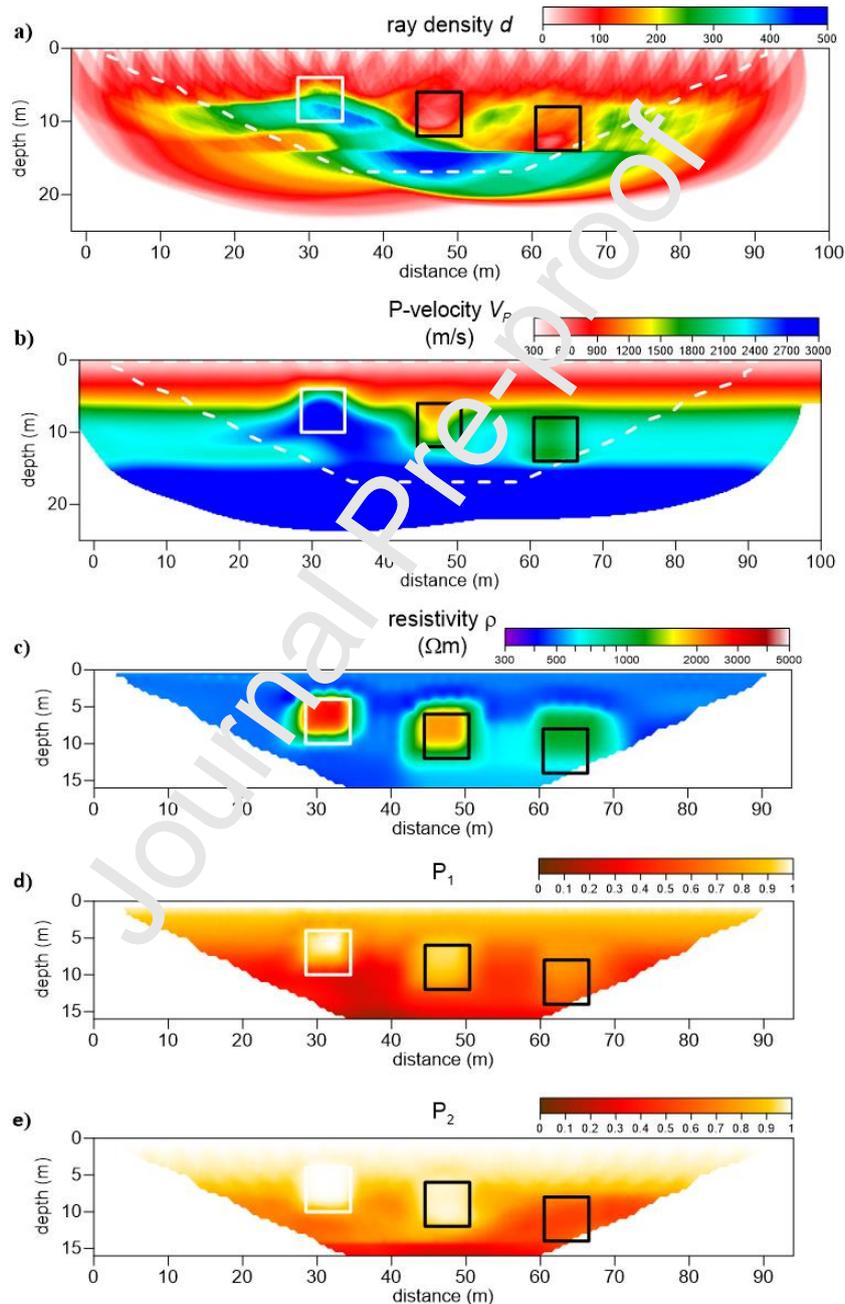


Fig. 7. Model C: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities and compact blocks are indicated by black and white solid lines respectively. The joint interpreted section is contoured by a dashed white line.

Even for this model the cluster analysis is dependent on the choice of the value of k classes (Fig. 8). The three anomalous zones are identified if $k > 2$ (Fig. 8 a), however only by choosing 5 clusters it is possible to discriminate the two cavities from the compact calcarenitic block that is identified by a specific cluster, while the two cavities are identified by another cluster (Fig. 8 d).

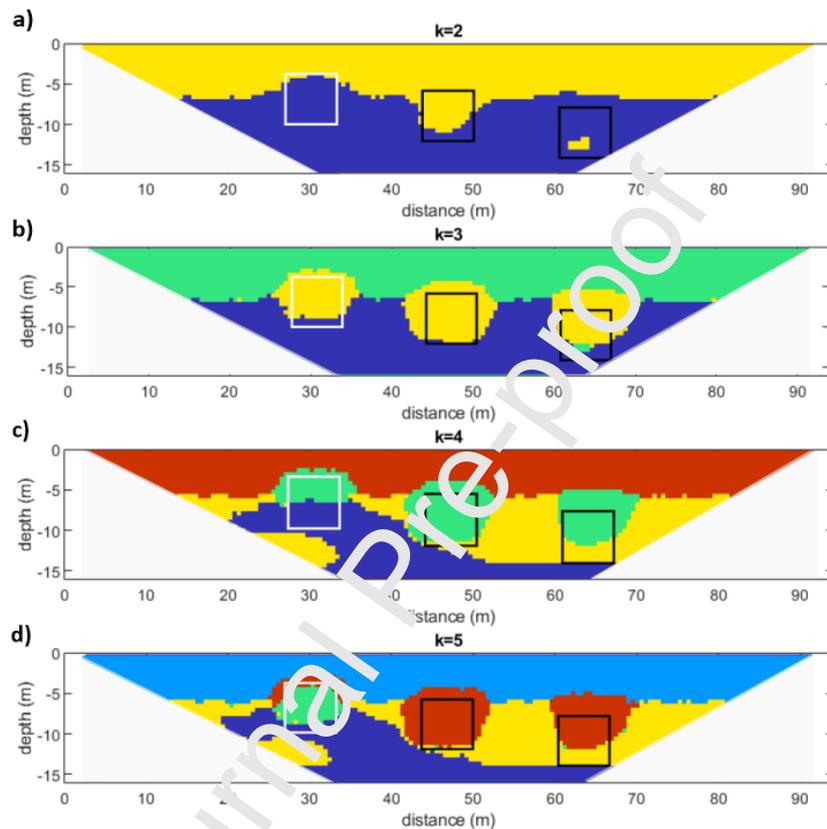


Fig. 8. Model C: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

3.4 Model D

In Fig. 9 the results of the inversion of simulated data obtained starting from the model D are showed. The inversion processing of the seismic data required 50 iterations to reach a misfit value of $12 \mu\text{s}$.

The distribution of the ray path density (Fig. 9 a) highlights two areas of low density in the approximate correspondence of the cavities; while a high density area is located at the center of the image, where the calcarenitic block is located, starting from a depth of 6 m to about 18 m, and spreading below the two cavities, marking their lower limits. SRT (Fig. 9 b) detects the decrease in seismic velocity in the two cavities, while maximum velocity values are at the center of the image corresponding to the position of the calcarenitic block. ERT (Fig. 9 c) well identifies the most

superficial cavity at 4 m depth with the highest resistivity values; but it associates low values of resistivity and very smoothed anomalies both to the cavity at 8 m of depth and to the calcarenitic block at 6 m of depth.

The likelihood parameter P_1 (Fig. 9 d) shows values close to 1 corresponding the shallower cavity (4 m of depth), while similar values, between 0.5 and 0.7 are showed by the cavity at 8 m depth and the calcarenitic block at 6 m depth. The likelihood parameter P_2 (Fig. 9 e) shows values close to one for the cavity at 4 m depth and very high for that at 8 m depth. Vice versa, the calcarenitic block is characterized by high parameter values, but a heterogeneous pattern that does not allow a clear identification of its shape.

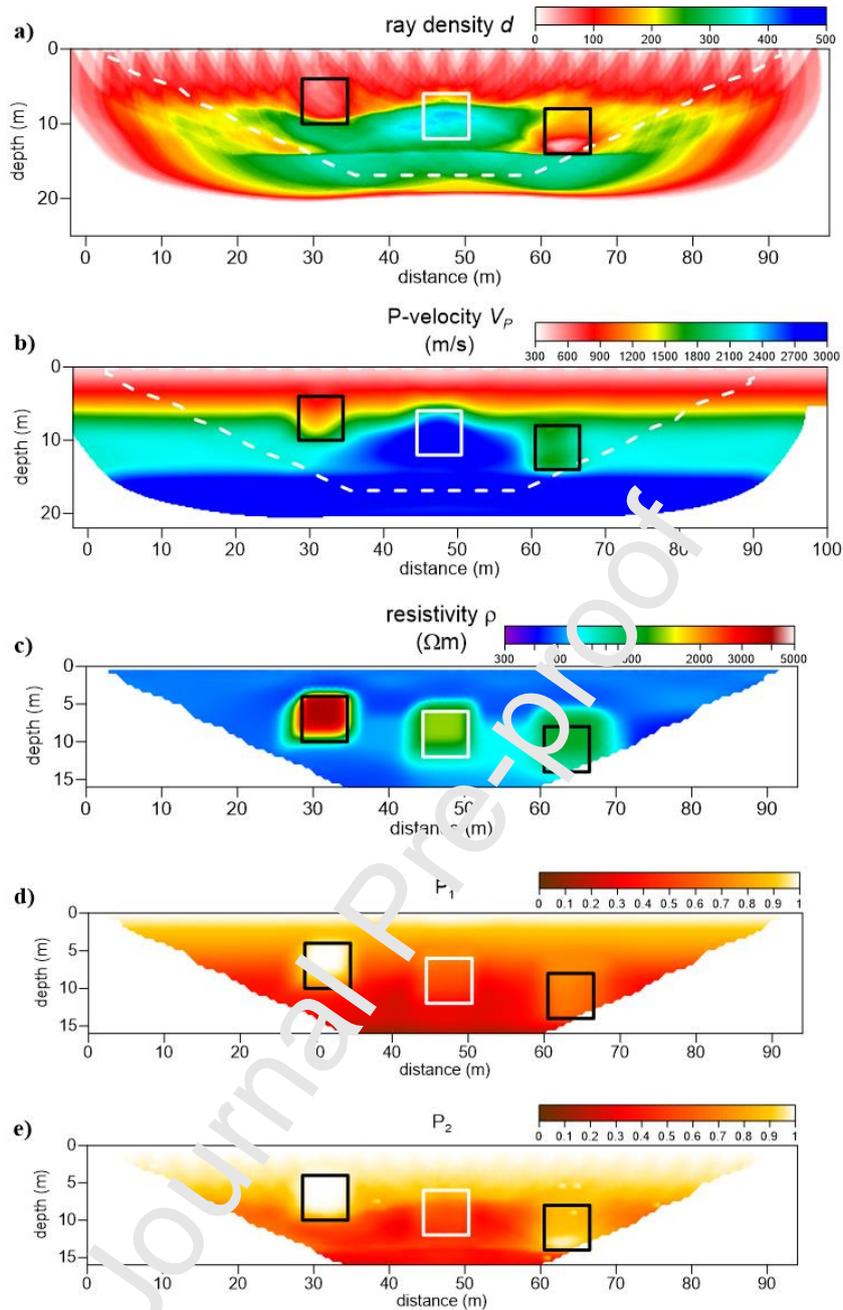


Fig. 9. Model D: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities and compact blocks are indicated by black and white solid lines respectively. The joint interpreted section is contoured by a dashed white line.

Images of spatial distribution of clusters for this model (Fig. 10) are less clear than those relating to the previously discussed model: the three anomalous zones are identified starting from $k \geq 4$ but, even for this model, the cavities and the calcarenitic block are discriminated by two different clusters only if we chose $k = 5$ (Fig. 10 e).

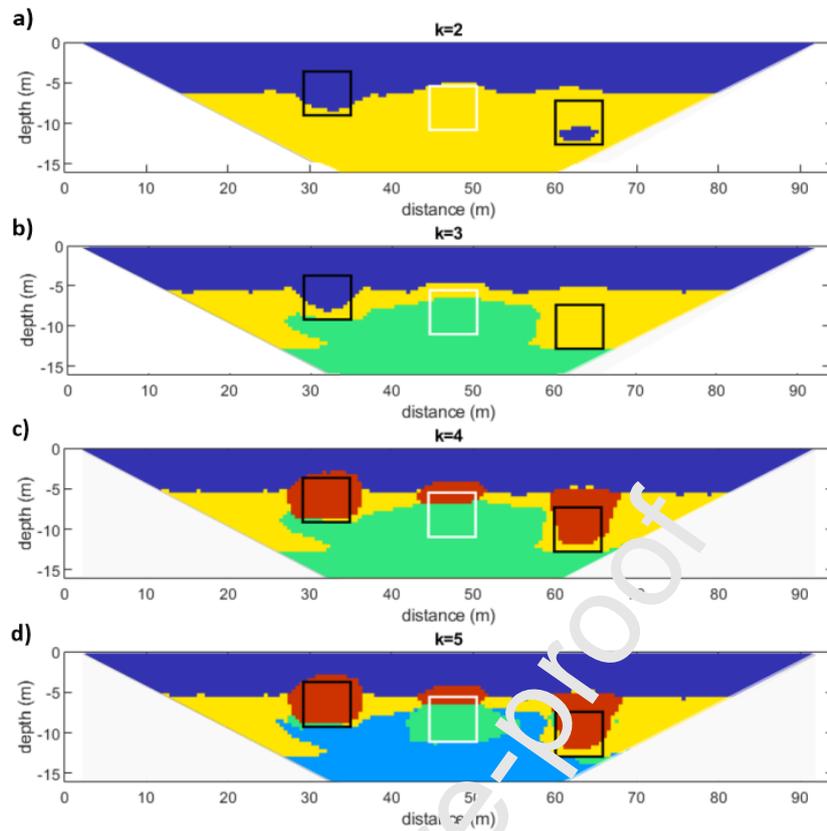


Fig. 10. Model D: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

3.5 Model E

The results of the inversion of simulated data obtained starting from the model E are showed in Fig. 11. After 50 iterations a misfit of 11 μs was obtained between predicted and measured travel times of P-waves. Looking at the distribution of the ray path density (Fig. 11 a) it's easy to identify two areas of low density at the cavities at 4 m and 6 m depth; while a high density zone is found in correspondence of the calcarenitic block, at a depth of 8 m. this area is not limited to the calcarenitic block but extends deep beneath the base of the cavities up to 20 m depth. SRT (Fig. 11 b) shows high velocity values that are located at the calcarenitic compact block but are not limited by this, extending laterally outside it. Non-significant decreases of P-wave velocity are found at the cavities while low values are instead found in the cavity areas at 4 m and 6 m depth. ERT (Fig. 11 c) discriminates quite well with high resistivity values the two surface cavities at 4 m and 6 m of depth, while to the calcarenitic compact block it associates a highly smoothed area with low values of resistivity. Both parameters P_1 and P_2 , but a little better P_2 , discriminate well the cavities showing values close to one in their correspondence. On the other hand, the compact calcarenitic block is poorly highlighted (Fig. 11 d, e).

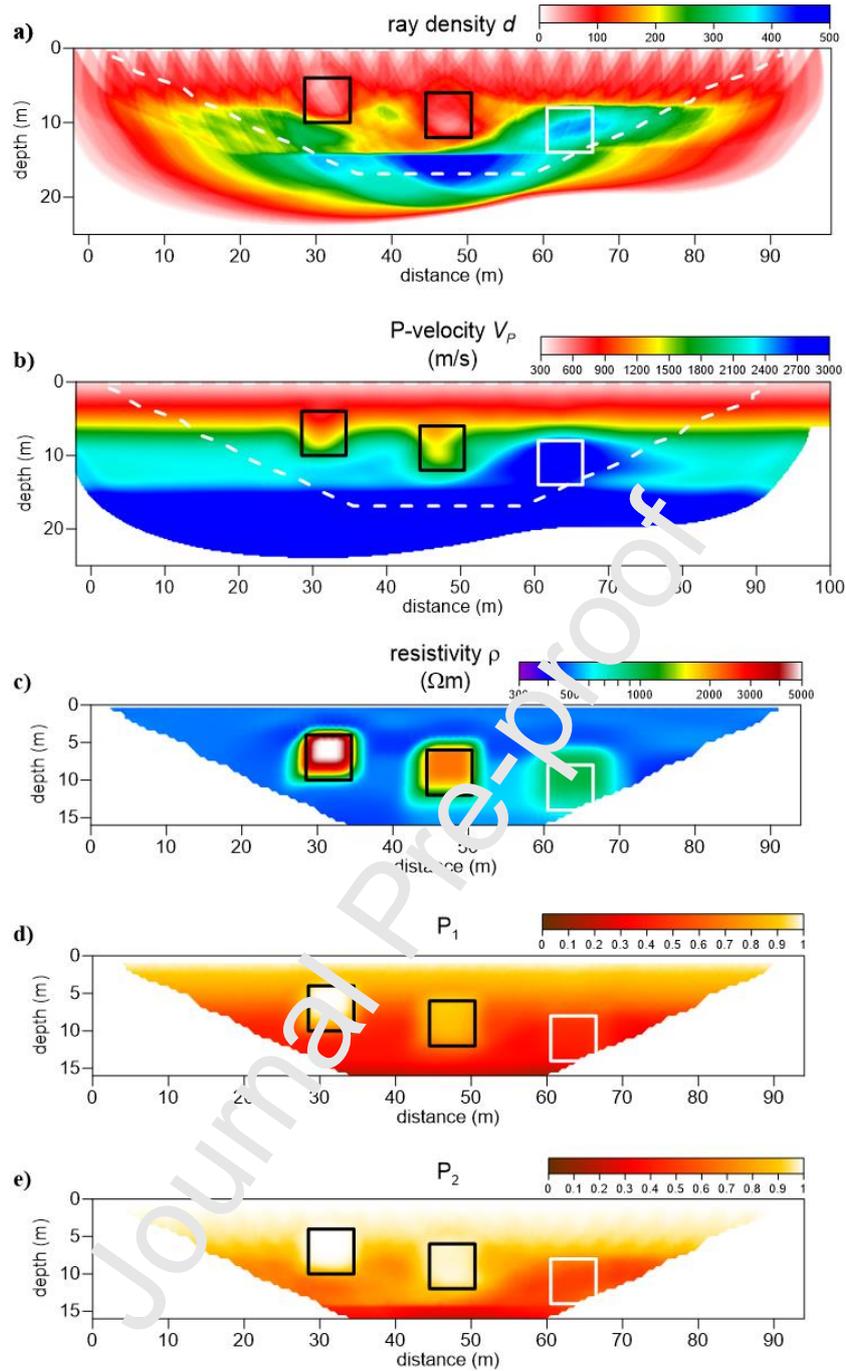


Fig. 11. Model E: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities and compact blocks are indicated by black and white solid lines respectively. The joint interpreted section is contoured by a dashed white line.

Starting from $k = 3$ the cluster analysis (Fig. 12) includes the two cavities in a specific cluster and well delimits their shape. Instead, the calcarenitic block, due to its depth, is never clearly identified, even if for $k \geq 4$ the images show a cluster that encloses it but that also extends in depth to areas of high ray path.

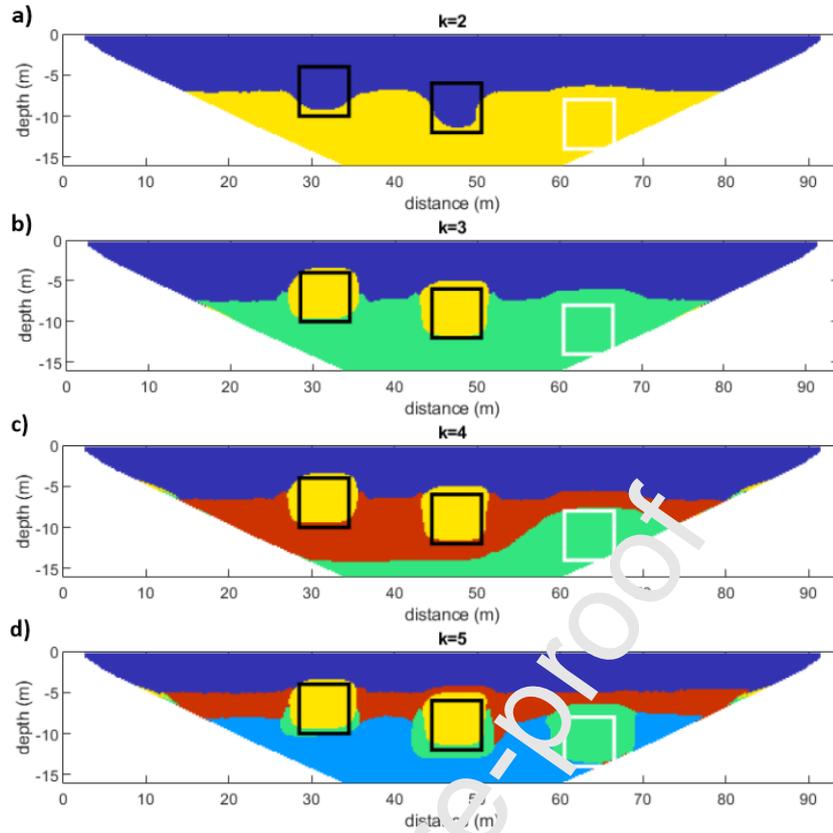


Fig. 12. Model E: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

3.6 Model F

Figure 13 shows the results of the inversion of simulated data obtained starting from the model F. The inversion processing of the seismic data required 20 iterations to reach a misfit value of 25 μ s. Considering the distribution of the ray path density (Fig. 13 a), three zones of low density, located from 28 m to 34 m, from 45 m to 50 m, from 58 m to 65 m with respect to the origin, correspond to the three cavity zones all located at a depth of 6 m. The fourth cavity extends also beyond the coverage area, and therefore cannot be fully investigated, nevertheless it also presents for the portion investigated a decrease, smaller than the other cavities, of the ray path density. An area of high ray coverage is located below the cavities between 14 m depth and 23 m depth, which perfectly delimits the base of the voids. SRT (Fig. 13 b) shows seismic velocity decreases in correspondence to all the cavities. ERT (Fig. 13 c) highlights cavities with four areas of high resistivity (values higher than 5000 Ω m are in correspondence with the greatest cavity). The distribution of likelihood parameters P_1 and P_2 (Fig. 13 d, e) shows a general distribution of values close to 1 both superficially than within the cavities. In this case the bottom and the lateral boundaries of the cavities are delimited quite well, but not the top.

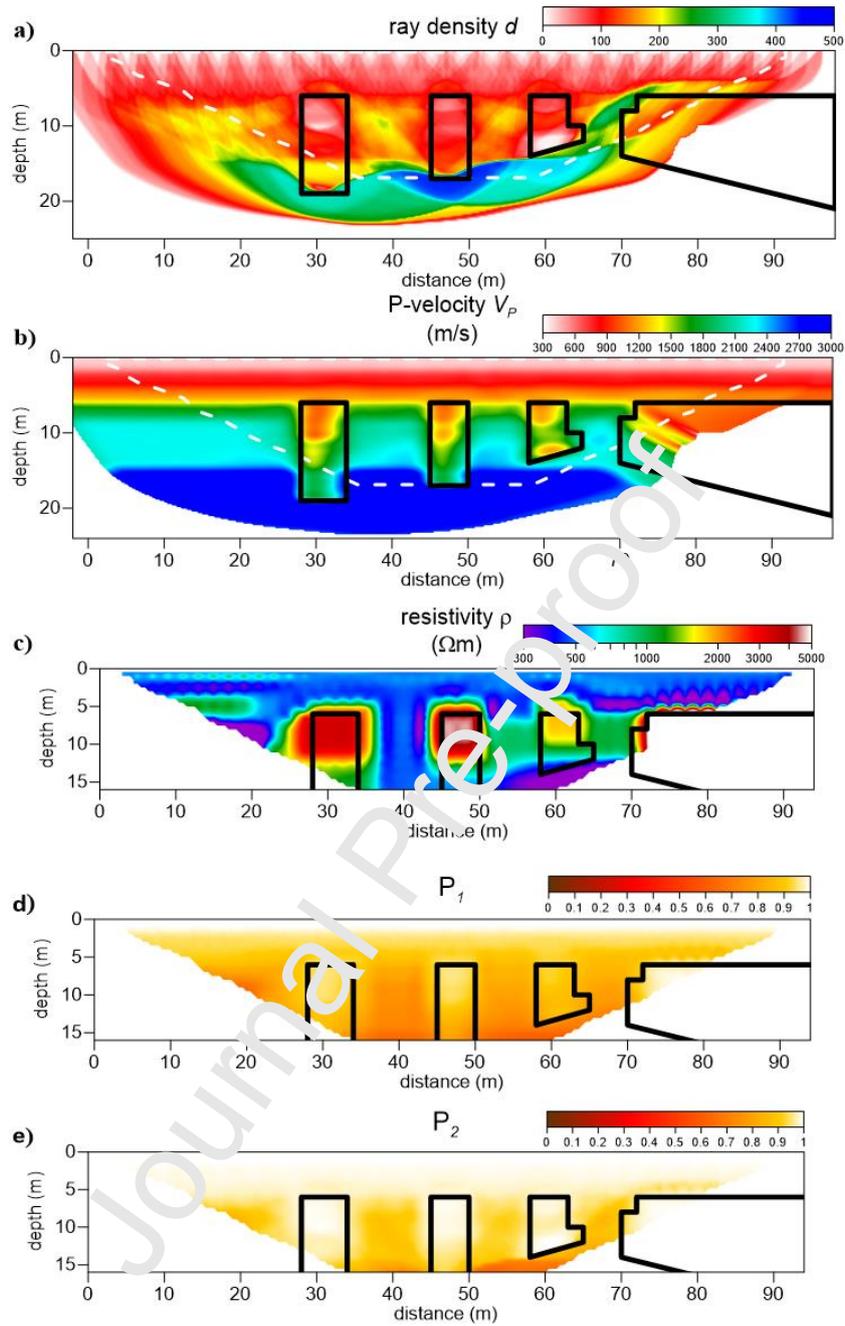


Fig. 13. Model F: results of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SRT of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . Boundaries of cavities are indicated by black solid lines. The joint interpreted section is contoured by a dashed white line.

Depending on the choice of the value of k classes, the cluster analysis returned different representation images. With $k = 2$ (Fig. 14 a) the image identifies two different clusters, almost including the cavities in a single cluster that comprises also the shallower subsoil. With $k = 3$ (Fig. 14 b) the cavities are specifically represented by a separate cluster even if its shape do not extend to the real depths of the cavities. With $k = 4$ and $k = 5$ (Fig. 14 c, d) the deep extension of the cavities

is not yet well identified by the cluster, however cluster analysis generates a new cluster that extends beneath the cavities delimiting their bottoms.

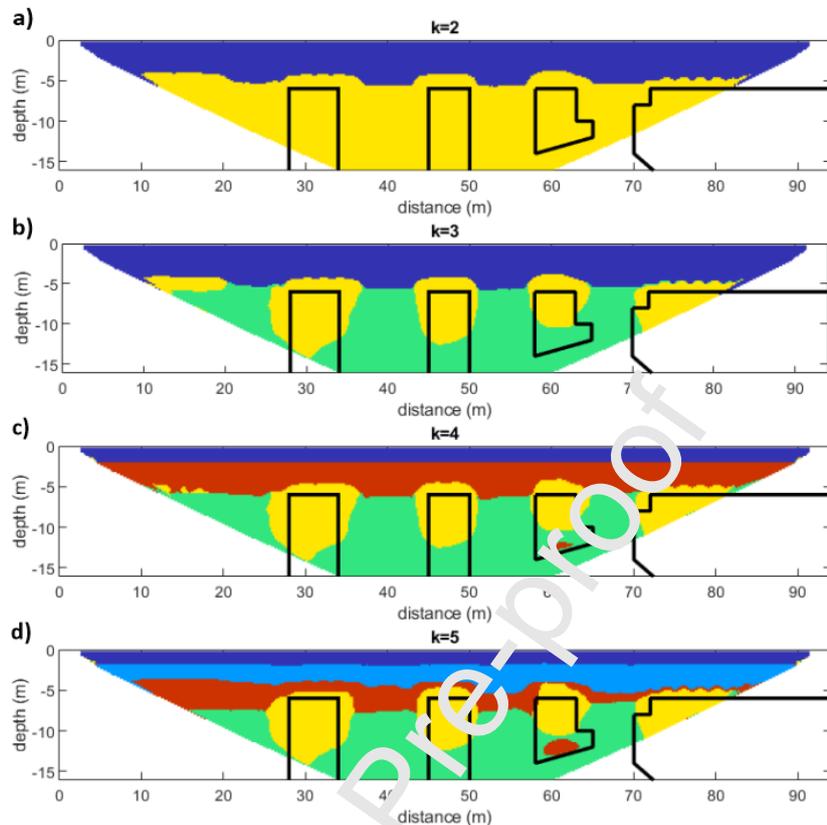


Fig. 14. Model F: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

4. Test site data analysis

Experimental data were acquired at the test site of Parco delle Cave of Marsala (Italy, Sicily) to compare theoretical and experimental data, optimize the inversion process and simplify the step of interpretation of the same data.

Since the time of the Phoenicians up to the present day this area has been exploited for the extraction of the Marsala calcarenite (Pleistocene Medio Inf.). This calcarenite is rather homogeneous and generally well-cemented, light yellow in color, poorly fossiliferous, with high thicknesses, with calcareous granules and carbonate cement. The rock is vacuolar, highly porous and poorly fractured.

The Foderà quarry (Fig. 2), recently defined also as “Parco delle Cave”, is now used for touristic and cultural activities.

The quarries of this area are mainly open-air quarries, with almost regular shapes and sub-vertical excavation fronts. Most of these quarries have now been filled, even if partially, with waste materials coming from excavation faces of the same quarry or nearby cavities.

In the Park there are also several underground cavities, located at varying depths with respect to the ground level, made by excavation of tunnels sustained by pillars. In more recent times, particularly in the post-war period, the latter were mostly abandoned due to various causes, as the huge cost of excavating the material, or the chemical-physical degradation of the calcarenite that generates instability phenomena like collapses or deformations of the pillars and vaults (Bonamini et al., 2013). This situation of abandonment led to an increase in the exploitation of open-air quarries, allowing to intercept the hypogeal areas.

The method of open excavation involves the removal of the surface alteration blanket that generally presents discontinuity surfaces, infiltration zones or physical alterations due to the long exposure from exogenous agents. And it allows to exploit the underlying rock with more appropriate characteristics for building purposes. The cultivation in underground instead, a technique currently in disuse, is practiced in order not to damage the upper ground. Access to the tunnels is through wells that are also used for lighting, ventilation of the tunnels and exit routes for the excavated sections. The progress procedure takes place with the formation of underground tunnels and halls of significant size, paying particular attention to leaving massive pillars supporting the rocky vaults. This technique is particularly decisive when the lithotype is of poor quality and therefore difficult to use.

The underground quarry chosen as field test develops in a series of tunnels and connected rooms, most of which have been the subject of a phase of hypogeographical topographic surveys with tools and software used for speleological activities. In particular, the Leica X310 laser distance gauge was used with precision in measurements of 2 mm and 0.5° RMS for angular measurements. The calcarenitic cavities of the Park of the Caves are almost all accessible (there are at least 3 entrances), in fact the relief carried out has allowed to determine the various configurations of the voids, the depths and their dimensions, allowing to rebuild the plant of the area and the reference section (Fig. 2 c).

The refraction seismic tomography investigation was carried out using 48 geophones with 2 m intervals between the geophones. Along the same alignment in the NE-SW direction, an electrical resistivity tomography was acquired with 48 electrodes, putting the electrodes in the same positions as the geophones. The inversion of the seismic data (after 50 iterations and a misfit value of 0.012 ms) has allowed to obtain the seismic rays density d and of P-waves velocity v_p . The results of the inversion are showed in Fig. 15. These distribution maps are characterized by a low seismic ray density and low P-wave velocity values in correspondence with the cavity blocks (Fig. 15 a, b). Also, the electrical resistivity model (Fig. 15 c) identifies three of the four cavities present in the real model. The two proposed correlation parameters are able to differentiate with greater precision

the cavity (values close to 0) from the high compact rock (high values), except for the cavity positioned at the end of acquisition line, due to the low resolution of the data in this area (Fig. 15 d, e).

It must be said that the complexity of the shape of the tunnels and chambers of the Foderà Quarry implies that the 2D inversions are affected by the ghost anomalies generated by the three-dimensionality of the intercepted structures and by the presence of other structures close to the survey lines (Martorana et al, 2018). Moreover, the 2D ERT carried out in situ is also disturbed by variations in the resistivity of the rock, and by the detritus, which are not foreseen by the theoretical model.

Journal Pre-proof

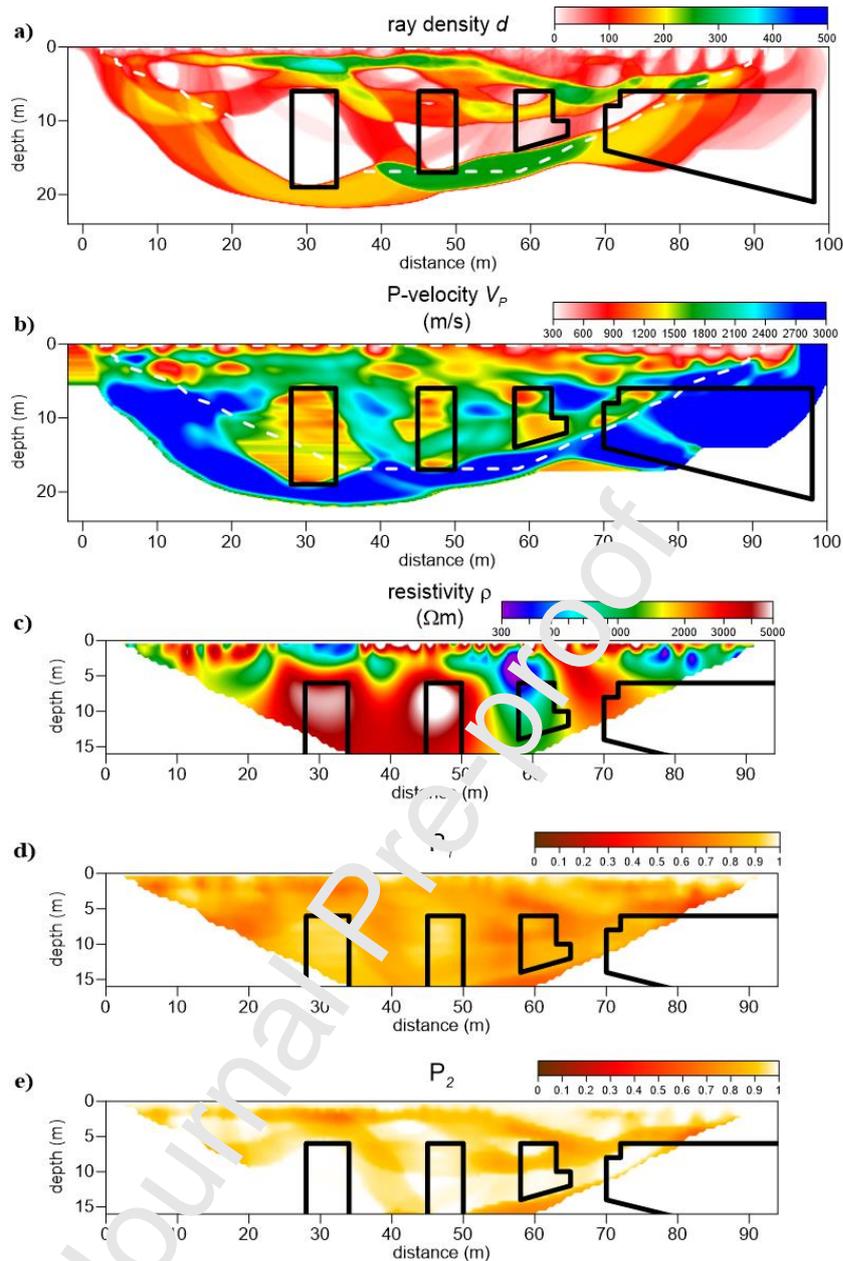


Fig. 15. Foderà Quarry: result of inversion of simulated P-waves travel-times and apparent resistivity data. a) ray path density d ; b) SR1 of P-waves velocity v_p ; c) ERT; d) imaging of the likelihood parameter P_1 ; e) imaging of the likelihood parameter P_2 . The joint interpreted section is contoured by a dashed white line.

Finally, the results of a non-hierarchical cluster analysis algorithm for a number of clusters from 2 to 5 are showed (Fig. 16). With $k = 2$ (Fig. 16 a) the map identifies two different clusters, grouping almost all cavities in a single cluster. From $k = 3$ to $k = 5$ (Fig. 16 b, c, d) the first two cavities are grouped in the same cluster, while the third cavity is associated to a different cluster.

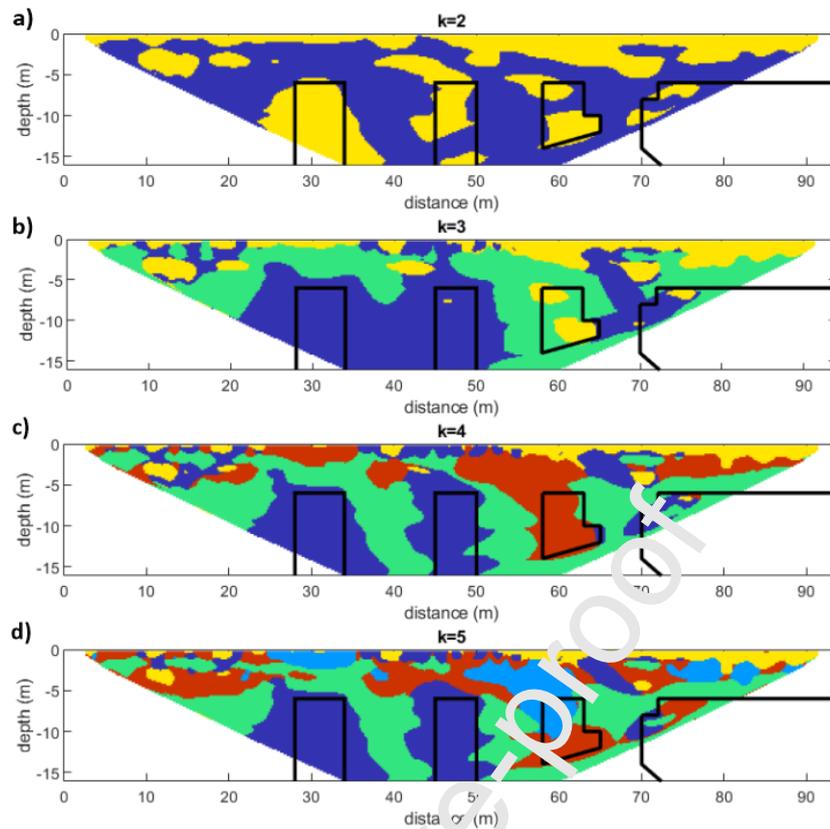


Fig. 16. Foderà Quarry: results of cluster analysis by choosing different number of classes k . a) $k = 2$; b) $k = 3$; c) $k = 4$; d) $k = 5$.

5. Discussion

Results of synthetic tests show that the detectability of the voids and the compact blocks of limestone is strongly influenced by the depth of the target and by its lateral position with respect to the length of the tomography. This obviously applies both considering the seismic velocity and the electrical resistivity as the investigation parameter. The electrical resistivity certainly gives higher contrasts than the seismic velocity, but it cannot allow to distinguish a cavity from a compact block of calcarenite with sufficient reliability, especially if this is close enough to the surface, to generate an evident anomaly. On the other hand, seismic refraction tomography, although generating anomalies of different sign (positive if in the presence of compact blocks, negative if in the presence of voids), does not accurately identify the geometries of the structures, due to the intrinsic limitations of the method.

The use of the likelihood parameters can help the interpretation, facilitating the distinction between the voids (in which the values of P_1 and P_2 tend to 1), from other structures that produce lower values, but also their yield depends on position and depth of the target. The P_2 shows contrasts greater than P_1 , if in the presence of voids, but a lower geometric precision.

The cluster analysis is able to give good results when choosing a sufficiently high number of k ($k \geq 4$ if it is necessary to discriminate between voids and compact blocks, also taking into account the stratification). However it must always be compared with the tomographic images to define better the shape of the structures.

The comparison between the results of filed survey and model F, based on the speleological survey, confirm the effectiveness of the proposed methods, but highlight other problems not considered in the synthetic models, with regard to the lithological heterogeneities and the three-dimensionality of the cavities that can determine for them assignment to a different cluster.

All cluster data relative to the cavity have been highlighted, for both synthetic data and experimental ones (yellow squares) (Fig. 17). To facilitate the comparison between these two graphs the values of the three components (seismic ray density, P-waves velocity and electrical resistivity) have been normalized. In this way it was possible to highlight how the cavities, considering the values of the three parameters analyzed, are arranged in the same region of the multiparameter space favoring a rapid identification of the cavities present in the subsoil. The region in the multiparametric space where the cavities are located is characterized by high values of electrical resistivity and low values of seismic ray density and P-waves velocity.

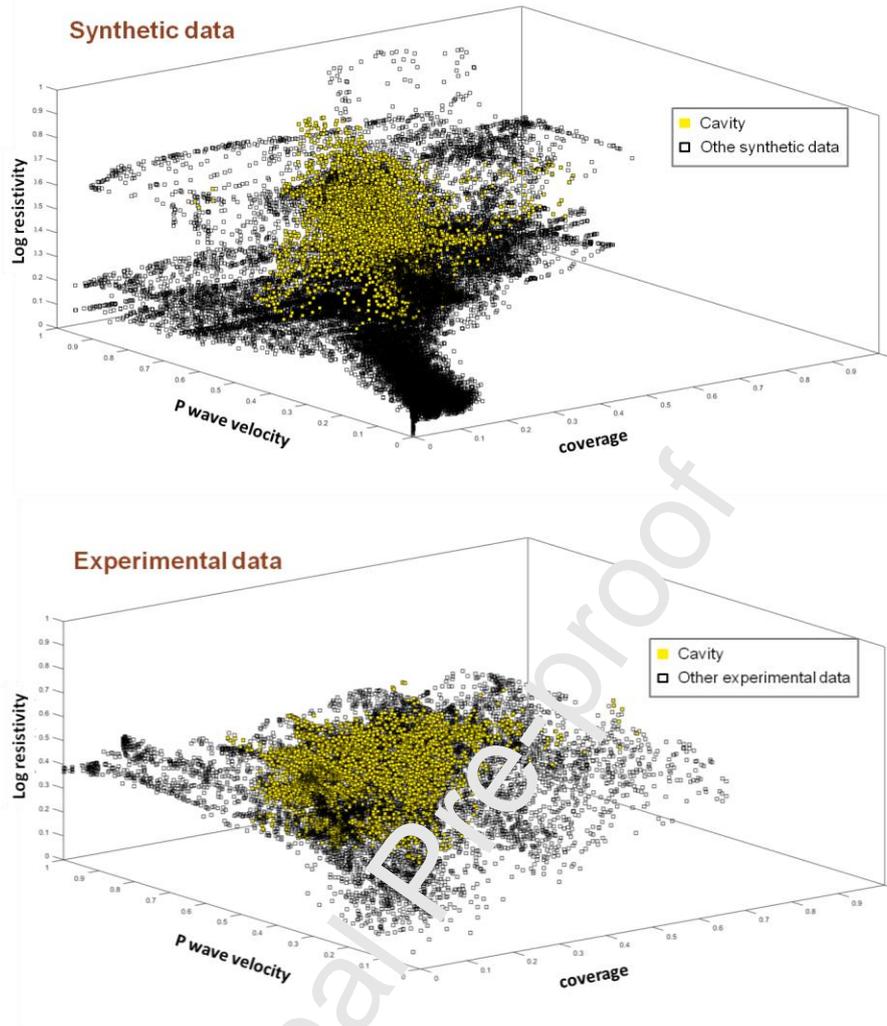


Fig. 17. Synthetic data (top) and experimental ones (bottom) related to the Foderà Quarry, plotted in the multiparameter space (seismic ray density, P-waves velocity and electrical resistivity). Yellow squares represent data relative to cavities.

6. Conclusions

This work shows how the joint study of seismic refraction tomography and electrical resistivity tomography techniques leads to the determination of cavities buried in the subsoil, minimizing possible interpretative ambiguities and producing the most robust results at the same time.

In order to facilitate discrimination between voids and other lithological structures, two likelihood parameters that correlate the physical quantities has been defined. The use of these parameters facilitates the distinction between voids and other structures but their effectiveness depends on the position and depth of the investigate target.

The cluster analysis performed on static units defined by electrical resistivity values, P-wave velocities, and seismic density on coincident sections, seems to be an effective method to characterize structures near the surface in the presence of complex geological settings. The use of

the non-hierarchical clustering algorithm has been chosen because it is less influenced by abnormal values, and allows a statistical unit to change its cluster during the iterative process. Depending on the choice of the number of clusters to be identified (values between 2 and 5), cluster distribution maps have been constructed in the multi-parameter space, allowing to define certain variability limits for the selected parameters, for synthetic and experimental data. Finally, experimental data show that electrical and seismic tomographies is not influenced by the presence of cavities without lateral continuity.

References

- Archie, G.E., 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. I. *Pet. Technol.* 5, doi: <http://dx.doi.org/10.2118/942054-G>.
- Barbarito, L., 1999. *L'analisi di settore: metodologia e applicazioni*, Milano, Franco Angeli Ed.
- Bonamini, M., Di Maggio, C., Lollino, P., Madonia, G., Parisi, M., Vattano, M., 2013. Study of anthropogenic sinkholes in the Marsala area (western Sicily) through field surveys and numerical analyses of instability processes in underground quarries. In *Geotane 2011, VIII Forum Italiano di Scienze della Terra* (pp. 74-74). IT, <http://hdl.handle.net/10447/60772>.
- Bottari, C., Albano, M., Capizzi, P., D'Alessandro, A., Doumaz, F., Martorana, R., Moro, M., Saroli, M., 2018a. Recognition of earthquake-induced damage in the Abakainon necropolis (NE Sicily): results from geomorphological, geophysical and numerical analyses. *Pure and Applied Geophysics*, 175, 1, 133-148, doi: <https://doi.org/10.1007/s00024-017-1653-4>.
- Bottari, C., Capizzi, P., Cavallaro, D., Cerretti, M., D'Alessandro, A., Lodato, L., Martorana, R., Pisciotta, A., Scudero, S., 2018b. Coseismic damage at an archaeological site in Sicily, Italy: evidence of Roman earthquake surface faulting. *Surveys in Geophysics*, 39, 6, 1263-1284. doi: <https://doi.org/10.1007/s10712-018-9482-2>.
- Capizzi, P., Cosentino, P.L., Frandaca, G., Martorana R., Messina, P., 2005. 2D GPR and geoelectrical modelling: tests on man-made tunnels and cavities. 11th European Meeting of Environmental and Engineering Geophysics of the Near Surface Geoscience Division of the EAGE, Near Surface 2005; Palermo; Code 104313, B036 1-4, ISBN 90-737-8191-4, doi: <https://doi.org/10.3997/2214-4609-pdb.13.B036>.
- Cardarelli, E., Fischanger, F., Piro, S., 2008. Integrated geophysical survey to detect buried structures for archaeological prospecting. A case-history at Sabine Necropolis (Rome, Italy). *Near Surface Geophysics* 6(1), 15–20, doi: <https://doi.org/10.3997/1873-0604.2007027>.
- Cardarelli, E., Cercato, M., Cerreto, A., Di Filippo, G., 2010. Electrical resistivity and seismic refraction tomography to detect buried cavities. *Geophysical Prospecting*, 58(4), 685-695, doi: <https://doi.org/10.1111/j.1365-2478.2009.00854.x>.

- Dahlin, T., Bjelm, L., Svensson, C., 1999. Use of electrical imaging in site investigations for a railway tunnel through the Hallandsas Horst, Sweden. *Q. J. Eng. Geol. Hydrogeol.* 32, 163–172, doi: <https://doi.org/10.1144/GSL.QJEG.1999.032.P2.06>.
- Dell'Aversana, P., 2001. Interpretation of seismic, MT and gravity data in a thrust belt interpretation. *First Break* 19(6), 335–341, doi: <https://doi.org/10.1046/j.1365-2397.2001.00158.x>.
- 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. *J. Appl. Geophys.* 101, 108–115, doi: <https://doi.org/10.1016/j.jappgeo.2013.12.004>.
- Fabbris, L., 1983. *Analisi esplorativa di dati multidimensionali*, Cleup ed.
- Fasani, G.B., Bozzano, F., Cardarelli, E., Cercato, M., 2013. Underground cavity investigation within the city of Rome (Italy): A multi-disciplinary approach combining geological and geophysical data. *Engineering Geology* 152, 109–121, doi: <https://doi.org/10.1016/j.enggeo.2012.10.006>.
- Gallardo, L.A., Meju, M.A., 2003. Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophysical Research Letters*, 30(13), doi: <https://doi.org/10.1029/2003GL017370>.
- Gallardo, L.A., Meju, M.A., 2004. Joint two-dimensional DC Resistivity and Seismic travel time inversion with cross-gradients constraints. *Journal of Geophysical Research*, 109(B3), doi: <https://doi.org/10.1029/2003JB002716>.
- Gallardo, L.A., Meju, M.A., 2007. Joint two-dimensional cross-gradient imaging of seismic travel time data for structural and lithological classification. *Geophysical Journal International*, 169(3), 1261–1272, doi: <https://doi.org/10.1111/j.1365-246X.2007.03366.x>.
- Gallardo, L.A., Meju, M.A., 2011. Structure coupled multiphysics imaging in geophysical sciences. *Rev. Geophys.* 49, RG1003. <http://dx.doi.org/10.1029/2010RG000330>.
- Gebrande, H., Miller, H., 1987. *Refraktionsseismik Angewandte Geowissenschaften II* ed F Bender (Stuttgart: F Enke) pp 220–60.
- Gibson, B.S., Odegard, M.E., Sutton, G.H., 1979. Nonlinear least-squares inversion of traveltimes data for a linear velocity-depth relationship. *Geophysics*, 44(2), 185–194, doi: <http://dx.doi.org/10.1190/1.1440960>.
- Hartigan, J. A., 1975. *Clustering algorithms*. John Wiley & Sons, Inc.
- Hellman, K., Ronczka, M., Günther, T., Wennermark, M., Rücker, C., Dahlin, T., 2017. Structurally coupled inversion of ERT and refraction seismic data combined with cluster-based model integration. *Journal of Applied Geophysics*, 143, 169–181, doi: <https://doi.org/10.1016/j.jappgeo.2017.06.008>.
- Imposa, S., Grassi, S., Di Raimondo, S., Patti, G., Lombardo, G., Panzera, F., 2018. Seismic refraction tomography surveys as a method for voids detection: an application to the archaeological park of Cava Ispica, Sicily, Italy. *International Journal of Architectural Heritage*, 12(5), 806–815. doi: [10.1080/15583058.2017.1419311](https://doi.org/10.1080/15583058.2017.1419311).

- Kotyrba, B., Schmidt, V., 2014. Combination of seismic and resistivity tomography for the detection of abandoned mine workings in Münster/Westfalen, Germany: Improved data interpretation by cluster analysis. *Near Surface Geophysics*, 2014, 12(3), 415-425. doi: <https://doi.org/10.3997/1873-0604.2013056>.
- Lees, J.M., VanDecar, J.C., 1991. Seismic tomography constrained by Bouguer gravity anomalies: applications in western Washington. *Pure Appl. Geophys.* 135, 31–52, doi: <https://doi.org/10.1007/BF00877007>.
- Linder, S., Paasche, H., Tronicke, J., Niederleithinger, E., Vienken, T., 2010. Zonal cooperative inversion of crosshole P-wave, S-wave, and georadar traveltime data sets. *J. Appl. Geophys.* 72 (4), 254–262, doi: <https://doi.org/10.1016/j.jappgeo.2010.10.003>.
- Loke, M.H., 2013. Tutorial: 2-D and 3-D Electrical Imaging Surveys. www.geotomosoft.com.
- Maraio, S., & Bruno, P.P.G., 2015. Near-surface Voids in the Neapolitan Volcanic Tuff (Italy) Detected by Seismic Refraction Tomography. In *Near Surface Geoscience 2015-21st European Meeting of Environmental and Engineering Geophysics* (Vol. 2015, No. 1, pp. 1-5). European Association of Geoscientists & Engineers, doi: <https://doi.org/10.3997/2214-4609.201413748>.
- Martorana, R., Capizzi, P., Carollo, A., 2018. Misinterpretation caused by 3D effects on 2D Electrical Resistivity Tomography: tests on simple models. In *2nd Conference on Geophysics for Mineral Exploration and Mining* (Vol. 2018, No. 1,) pp. 549-566). European Association of Geoscientists & Engineers, doi: <https://doi.org/10.3997/2214-4609.201802560>.
- Martorana, R., Capizzi, P., D'Alessandro, A., Luzio, D., 2017a. Comparison of different sets of array configurations for multichannel 2D ERT acquisition. *J Appl Geophys*, 137, 34–48. doi: <https://doi.org/10.1016/j.jappgeo.2016.12.012>. ISSN: 0926-9851.
- Martorana, R., Capizzi, P., Avellone, G., D'Alessandro, A., Siragusa, R., & Luzio, D., 2017b. Assessment of a geological model by surface wave analyses. *Journal of Geophysics and Engineering*, 14(1), 159-172, doi: <https://doi.org/10.1088/1742-2140/14/1/159>.
- McDonald, R., Davies, R., 2003. Integrated geophysical surveys applied to karstic studies. *First Break*, 21(10).
- Meju, M.A., Gallardo, L.A., Mohamed, A.K., 2003. Evidence for correlation of electrical resistivity and seismic velocity in heterogeneous near-surface materials. *Geophys. Res. Lett.* 30, 1373. <http://dx.doi.org/10.1029/2002GL016048>.
- Orfanos, C., Apostolopoulos, G., Amolochitus, G., Leontarakis, K., 2008. Integrated geophysical approach for the detection of underground voids in a Construction site. 70th EAGE Conference & Exhibition, Rome, Italy, Extended Abstracts, P097.
- Orfanos, C., Apostolopoulos, G., 2013. Multiparameter analysis of geophysical methods for target detection: The unified geophysical model approach. *Geophysics*, VOL. 78 no. 6, P. IM1–IM13, doi: <https://doi.org/10.1190/geo2012-0285.1>.

- Qin F., Luo Y., Olsen K., Cai W., Schuster G.T., 1992. Finite-difference solution of the eikonal equation along expanding wavefronts. *Geophysics*, 57(3), 478-487, doi: <http://dx.doi.org/10.1190/1.1443263>.
- Rayfract 2010a: Palmer 2010 Syncline Model construction and forward modeling, with Rayfract® 3.25 and Golden Software Surfer® 11. Downloaded from <http://rayfract.com/tutorials/palmfig9.pdf>.
- Rayfract 2010b: Basement Step Mode construction and forward modeling, with Rayfract® 3.25 and Golden Software Surfer® 11. Downloaded from <http://rayfract.com/tutorials/step.pdf>.
- Riddle, G.I., Hickey, C.J., Schmitt, D.R., 2010. Subsurface tunnel detection using Electrical Resistivity Tomography and Seismic Refraction Tomography: a case study. In 23rd EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems (pp. cp-175). European Association of Geoscientists & Engineers, 552-562, doi: <https://doi.org/10.3997/2214-4609-pdb.175.SAGEEP064>.
- Rohdewald, S.R., 2011. Interpretation of first-arrival travel times with wavepath eikonal travelttime inversion and wavefront refraction method. Symposium on the Application of Geophysics to Engineering and Environmental Problems. Society of Exploration Geophysicists, doi: <https://doi.org/10.4133/1.3614086>.
- Rohdewald, S.R., 2016. Rayfract Online Help [Online] Intelligent Resource Inc. Available: <http://www.rayfract.com>.
- Schuster, G.T., Quintus-Bosz, A., 1993. Wavepath eikonal travelttime inversion: theory. *Geophysics*, 58, 1314–23, doi: <http://dx.doi.org/10.1190/1.1443514>.
- Scudero, S., Martorana, R., Capizzi, P., Pisciotta, A., D'Alessandro, A., Bottari, C., Di Stefano, G., 2018. Integrated Geophysical Investigations at the Greek Kamarina Site (Southern Sicily, Italy). *Surveys in Geophysics*, 39, 6, 1181-1200. doi: <https://doi.org/10.1007/s10712-018-9483-1>.
- Sheehan, J.R., Doll, W.E., Watson, D.C., Mandell, W.A., 2005. Application of Seismic Refraction Tomography to karst cavities. US Geological Survey Karst Interest Group Proceedings, Rapid City, South Dakota, 12–15 September 2005, 29-38.
- Tronicke, J., Holliger, K., Burrasini, 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>.
- Van Schoor, M., 2002. Detection of sinkholes using 2D electrical resistivity imaging. *Journal of Applied Geophysics*. 50, 393–399, doi: [https://doi.org/10.1016/S0926-9851\(02\)00166-0](https://doi.org/10.1016/S0926-9851(02)00166-0).
- Wolke, R., Schwetlick, H., 1988. Iteratively reweighted least squares algorithms, convergence analysis, and numerical comparisons. *SIAM Journal on Scientific and Statistical Computing*, 9(5), 907–921, doi: <https://doi.org/10.1137/0909062>.
- Wyllie, M.R.J., Gregory, A.R., Gardner, L.W., 1956. Elastic wave velocities in heterogeneous and porous media. *Geophysics* 21 (1), 41–70, doi: <https://doi.org/10.1190/1.1438217>.

Carollo Alessandra: Data Curation, Visualization, Investigation, Writing - Original Draft, Writing - Review & Editing.

Capizzi Patrizia: Conceptualization, Formal analysis, Software, Writing - Original Draft, Writing - Review & Editing.

Martorana Raffaele: Methodology, Validation, Formal analysis, Writing - Original Draft, Writing - Review & Editing.

Journal Pre-proof

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- An integrated study of SRT and ERT is proposed for both synthetic and real data.
- A statistical approach based on K-means cluster analysis has been applied.
- Multi-space cluster distribution maps helped to better detect the targets.

Journal Pre-proof