

Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements

Alain Rabaute, Béatrice Yven, Walter Chelini, Maria Zamora

▶ **To cite this version:**

Alain Rabaute, Béatrice Yven, Walter Chelini, Maria Zamora. Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements. Journal of Geophysical Research : Solid Earth, 2003, 108, p. 219-245. 10.1029/2001JB001436 . insu-03598423

HAL Id: insu-03598423

https://insu.hal.science/insu-03598423

Submitted on 6 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements

Alain Rabaute,[1] Béatrice Yven,[2] Walter Chelini,[3] and Maria Zamora[2]

Received 2 October 2001; revised 25 July 2002; accepted 18 October 2002; published 27 March 2003.

[1] The volcanic complex of Phlegrean Fields, located northwest of Naples (Italy), has been the site of deep geothermal exploration in the 1980s. Several wells were drilled by an Agip-Enel joint venture, with downhole continuous physical properties acquired in each well. The main purpose of this study is to map and describe the spatial variations of the volcanic deposits, in terms of their nature, thickness and in situ physical properties. We apply fuzzy *k* means classification of the well-logging data available in the wells cored in the San Vito area. Fuzzy set theory provides an approach that quantitatively assigns individuals to physically continuous classes. The optimal number of classes is found by minimization of three mathematical functions, thus reducing subjectivity. The resulting classifications are found (1) to reflect the main lithologies in the San Vito plain, (2) to provide more detail on the geological stratigraphy, allowing a more precise volcanic history, and (3) to differentiate between transition zones and interbedded well-defined deposits. In addition to discriminating between lithologies that have different physical properties, this study gives information on the degree of homogeneity of each lithologic unit, and the range of variation for the measured properties. Finally, using the physical classification of the deposits, we are able to propose a detailed two-dimensional compressional acoustic velocity structure of the studied area. *INDEX TERMS:* 0915 Exploration Geophysics: Downhole methods; 0910 Exploration Geophysics: Data processing; 1749 History of Geophysics: Volcanology, geochemistry, and petrology; 5102 Physical Properties of Rocks: Acoustic properties; *KEYWORDS:* Phlegrean fields, downhole measurements, fuzzy clustering, acoustic velocity, physical properties, stratigraphy

Citation: Rabaute, A., B. Yven, W. Chelini, and M. Zamora, Subsurface geophysics of the Phlegrean Fields: New insights from downhole measurements, *J. Geophys. Res.*, *108*(B3), 2171, doi:10.1029/2001JB001436, 2003.

1. Introduction

[2] The Phlegrean Fields is a quaternary volcanic and hydrothermal system, located 10 km northwest of Naples (Italy). A 12-km-wide caldera constitutes the main volcano-tectonic feature of the area, inside which the volcanic, seismic, and fumarolic activities are confined. The eruptive and unrest phenomena mechanisms of this densely inhabited area (affected by two recent episodes of strong ground uplift and seismic events in 1969–1972 and 1982–1984) have been thoroughly studied, and their volcanic hazard has been assessed. These studies have been largely supported by the Agip-Enel (Italian national petroleum and electricity companies) joint venture, which investigated the economic interest of geothermal activity in the Phlegrean Fields. Two main contributions arose from this geothermal research [*Agip*, 1987]: (1) the surface geological and geophysical surveys provided detailed geologic, geoelectric, gravimetric, and magnetometric maps and (2) the result of the exploration drilling helped to reconstruct the eruptive history and the lithostratigraphic sequence of the explored areas.

[3] However, although the geological structure of the subsurface of the Phlegrean Fields is well described, little is known of the geophysical properties of the volcanic deposits in depth [*Zamora et al.*, 1994; *Di Maio et al.*, 2000; *Yven*, 2001]. This is partly the reason why there are still difficulties and uncertainties in explaining and modeling the mechanical response accompanying the unrest episodes, the knowledge of the physical properties of the surface and subsurface formations of an active volcanic system being essential to define a detailed and representative geophysical model of it.

[4] Well-logging techniques are currently applied in sedimentary areas, where they provide fundamental information on rock and fluid properties. By cross correlation between boreholes, they allow define the structure of the studied area. As their use in volcanic environments is uncommon, the data obtained by Agip in the wells provides a unique opportunity to investigate the in situ physical properties of the main volcanic deposits.

[5] The present study focuses on the in situ well-logging physical properties acquired in the San Vito plain, located

[1]Laboratoire de Géologie, UMR 8538, École Normale Supérieure, Paris, France.
[2]Institut de Physique du Globe de Paris, CNRS-ESA 7046, Université Paris 7, France.
[3]Agip S.p.A, Formation Evaluation Departamento, San Donato Milanese, Italy.

Copyright 2003 by the American Geophysical Union.
0148-0227/03/2001JB001436$09.00

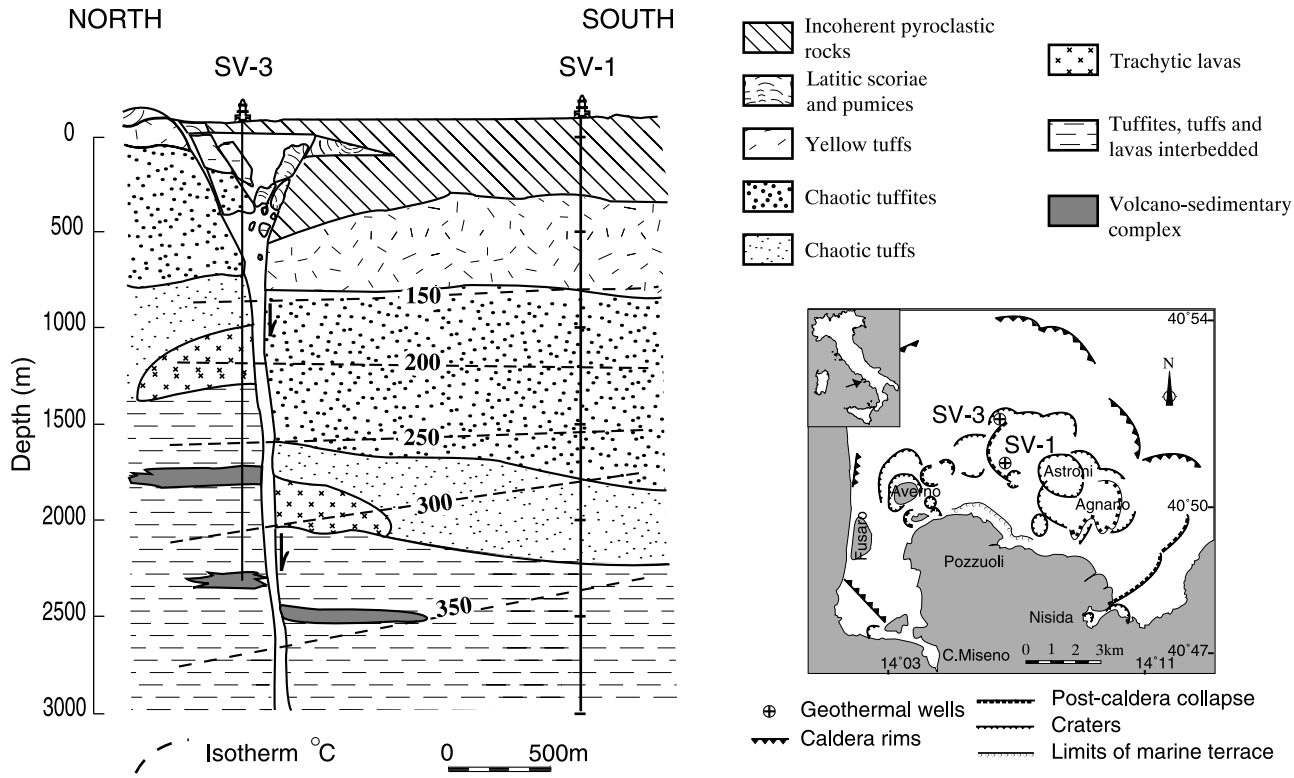


Figure 1. Lithostratigraphic cross section of the San Vito geothermal area (modified after *Rosi and Sbrana* [1987]).

3 km north of the city of Pozzuole (see Figure 1). The two studied wells, San Vito 1 (SV-1) and San Vito 3 (SV-3), cut 2360 and 3040 m, respectively of volcanic deposits and sedimentary rocks. They show a different geological history: SV-1 is located inside a collapse structure, due to the Gauro eruption area some 10,500 years ago, while SV-3 is outside. Furthermore, SV-1 is much closer than SV-3 to the center of the area where the recent ground deformations and seismic activity occurred. It is therefore interesting and useful to study their differences in physical properties. The downhole measurements are analyzed using fuzzy classification to give new insights on the different volcanic deposits in terms of their average in situ physical properties. As a result of this study, a detailed compressional acoustic wave velocities two-dimensional (2-D) structure of the San Vito area is proposed.

2. Geological Settings

[6] The Phlegrean Fields is an active volcanic area located inside the deep and large tectonic graben that forms the Campanian plain. The volcanic activity of the Phlegrean Fields began probably before 50,000 years B.P. [*Rosi and Sbrana*, 1987]. However, the first widespread emitted products are related to the impressive eruption of pyroclastic rocks, known as the Campanian Ignimbrite, which occurred around 37,000 years ago. The precaldera activity was at first submarine, with deposits of trachytic and latitic lavas interbedded with tuffs, tuffites, and Quaternary siltstones and sandstones. The activity became more and more subaerial and ended with the Campanian Ignimbrite episode, provoking the collapse of the caldera, which was then filled with water. Until 10,500 years B.P., the activity is again submarine and is characterized by chaotic tuffites deposits intercalated with thin beds of trachytic to latitic lavas. The activity culminates between 14,000 and 10,500 years B.P. with the eruption of the Neapolitan Yellow Tuffs, emitted from several eruptive centers confined inside the caldera [*Rittmann et al.*, 1950; *Rosi et al.*, 1983]. Several major collapses resulted from this episode. One of them appears clearly in the lithostratigraphy cross section between San Vito 1 and 3 wells (Figure 1). The last period is mostly subaerial and is divided into two main active phases separated by a long repose period of 3500 years. A detailed history of the volcanic activity of the Phlegrean Fields is given by *Rosi et al.* [1983], *Rosi and Sbrana* [1987], *Orsi et al.* [1996], and *Di Vito et al.* [1999].

[7] The regional stratigraphy of the Phlegrean Fields, based on petrography and on radiometric ages of at least 150 sections coming from the geothermal wells drilled in the Mofete and San Vito areas, has been mainly described by *Rosi et al.* [1983]. Figure 1 illustrates the stratigraphy of the San Vito area. The presence of chaotic tuffs layer are prevalent on the lava dome and lava flows in San Vito.

3. Downhole Measurements

[8] Three types of physical properties were recorded in the wells:

1. Electrical resistivity is measured at different lateral depths of investigation. The electrical resistivity measurements were obtained using the dual laterolog tool. The shallow Laterolog (LLs) measures the resistivity of the zone invaded by the mud, in case of a permeable formation. The

Table 1. Physical Properties Measured Downhole and Corresponding Intervals[a]

	SV-1	SV-3
Type of measurement	electrical-nuclear-acoustical	electrical-nuclear-acoustical
Physical properties	LLd, LLs-GR-DT	LLd, LLs-GR-DT
Intervals, m	160–2153	204–2350

[a]DT is compressional acoustic slowness, LLd and LLs are deep and shallow depths resistivity, and GR is total natural gamma radiation.

deep Laterolog (LLd) measures the resistivity of the noninvaded zone. When the formation is not permeable LLs and LLd are equivalent.

2. Compressional acoustic wave slowness (DT) is measured, its reciprocal being the velocity of the compressional wave.

3. Bulk natural gamma radioactivity (GR) is very sensitive to Th-, U-, and K-bearing minerals, such as K-feldspars, clay minerals such as illite or kaolinite but also micas (biotite) and heavy minerals (zircon, monazite, apatite).

[9] The intrinsic vertical resolution of the tools used for the measurements is directly dependent on the tool configuration. The theoretical value has been assessed through measurements in reference holes: it is 20–31 cm for the measure of the natural radioactivity, 61 cm for the different resistivity measurements, and around 1 m for the acoustic slowness.

[10] Details on logging techniques can be found in classical textbooks such as those by *Serra* [1984], *Hearst and Nelson* [1985] and *Ellis* [1987] and in the collection of papers published by *Hurst et al.* [1990, 1992] and *Harvey and Lovell* [1998].

[11] Because of the important depth reached, logging operations in each well were carried out through multiple runs, each of them covering a fraction of the borehole. Depth overlaps between successive runs in a single hole allowed successful splicing, in order to obtain the most complete data coverage. As the result of the geothermal gradient of the San Vito area (150°C/km), recording was often stopped before the bottom of the hole was reached. Table 1 shows the type of measurement, the physical properties measured for each borehole, and the initial and final depths.

[12] The variables used in the fuzzy classification are (1) the compressional acoustic wave slowness (DT), (2) the bulk natural gamma radioactivity (GR), (3) the logarithm of the deep resistivity, and (4) the ratio of the deep resistivity over the shallow resistivity (LLd/LLs), which informs on the mud invasion and, indirectly, on the permeability and the water saturation of the formation.

[13] When one was available, we checked on the repeat section for precision of the measurement. An example of such comparison is shown in Figure 2. We see that there is a fairly good agreement between the actual measurement and the repeat section. The accuracy of the measurement was checked through calibration patterns given at the end of each logging run record. All resistivity data were filtered using a low-pass filter in an attempt to remove high-frequency noise.

4. Fuzzy (Continuous) Classification

[14] Classification is a simple way of reducing a data set, composed of a large number of observations made on several variables, to a few number of classes (clusters) having more or less distinct properties. In our case, it is useful to discriminate between the different lithologies encountered in a borehole from the properties measured by logging techniques. It has been shown [*Lofts*, 1993; *Rabaute et al.*, 1997] that a classical clustering algorithm, such as the k-means algorithm [*McQueen*, 1967; *Forgy*, 1965], can be used to infer a log-derived lithostratigraphy, helping to overcome poor or no core recovery. However, in the case study presented here, the geological units suffered from strong diagenesis due to fluid circulation and high temperature and pressure conditions. These transformations affected the entire formation, but with different degree from place to place. The boundaries between the different units, and the behavior of the physical properties inside one unit, are therefore sought to be gradual instead of abrupt. Classification into mutually exclusive classes with rigidly defined boundaries, that is, "hard" or discontinuous classification, may be unsuitable here. Fuzzy set theory provides the continuous approach we need.

4.1. Hard Classification

[15] If one considers a set X of n data that one wants to partition into p discontinuous classes, the result will be a $n \times p$ matrix of memberships $M = (m_{ik})$, where $m_{ik} = 1$ if

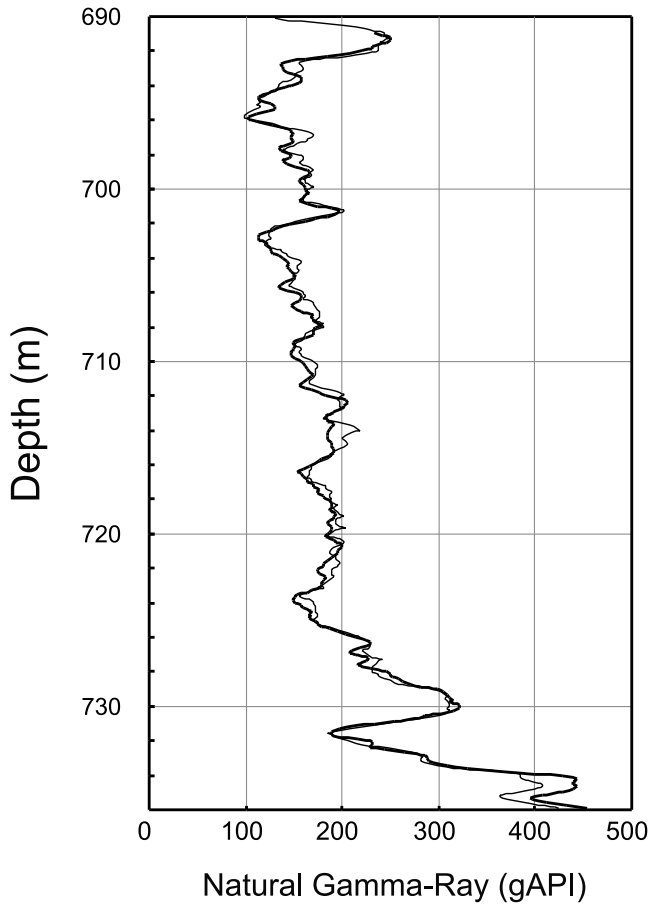


Figure 2. Repeat section of the total gamma ray curve. The thin curve is the main log; the thick curve is the repeat measurement.

data point i belongs to class k, and $m_{ik} = 0$ otherwise. The following conditions on M apply:

$$\sum_{k=1}^{p} m_{ik} = 1, \qquad i = 1 \ldots n \tag{1}$$

$$\sum_{i=1}^{n} m_{ik} > 0, \qquad k = 1 \ldots p \tag{2}$$

$$m_{ik} \in \{0, 1\}, \qquad i = 1 \ldots n; k = 1 \ldots p. \tag{3}$$

[16] The k-means algorithm minimizes the within-class sum-of-square errors function $J(M, C)$ under conditions 1, 2, and 3:

$$J(M, C) = \sum_{i=1}^{n} \sum_{k=1}^{p} m_{ik}\, d^2(x_i, c_k), \tag{4}$$

where $C = (c_{kv})$ is a $p \times q$ matrix of class centroids c_{kv} of class k for the variable v, with $v = 1 \ldots q$ and q, the number of variables; $x_i = (x_{i1}, \ldots, x_{iq})^T$ is the vector representing individual i, and $c_k = (c_{k1}, \ldots, c_{kq})^T$ is the vector representing the centroid of class k. $d^2(x_i, c_k)$ is the square distance between x_i and c_k according to a chosen definition of distance (noted d_{ik}^2 hereinafter).

4.2. Fuzzy Classification

[17] Starting with the works by *Zadeh* [1965], *Ruspini* [1969, 1970], and *Dunn* [1974], several methods for constructing continuous classes have been developed, the most popular being the fuzzy k-means method [e.g., *Bezdek*, 1981; *Bezdek et al.*, 1984]. Fuzzy k means is a direct generalization of hard k means [*Hartigan*, 1975], where the indicator function of conventional set theory, with value 0 or 1, is replaced by the membership function of fuzzy set theory, with values in the range 0 to 1.

[18] *Zadeh* [1965] and *Ruspini* [1969] developed the theory of fuzzy sets by relaxing the all-or-nothing status of the memberships so they are allowed to be partial, i.e., to take any value between and including 0 and 1. Condition 3 is then replaced by

$$m_{ik} \in [0, 1], \qquad i = 1 \ldots n; k = 1 \ldots p \tag{5}$$

leading to

$$J_F(M, C) = \sum_{i=1}^{n} \sum_{k=1}^{p} m_{ik}^2 d_{ik}^2. \tag{6}$$

Later, *Dunn* [1974] and *Bezdek* [1974] generalized equation (6) as

$$J_G(M, C) = \sum_{i=1}^{n} \sum_{k=1}^{p} m_{ik}^{\phi} d_{ik}^2, \tag{7}$$

where $\phi \in [1, \infty\}$ is the fuzzy exponent and controls the degree of fuzziness of the classification. With the smallest meaningful value $\phi \approx 1$, the solution of equation (7) is clearly a hard partition. Equation (7) is termed the fuzzy k-means objective function, and is minimized to satisfy conditions 1, 2, and 5 [*Bezdek*, 1981; *McBratney and de Gruijter*, 1992].

Table 2. Centroid and Standard Deviation of the Classes Resulting From the Fuzzy k-Means Analysis of Data Sets 1 and 2[a]

k	c_{kDT}, μs ft^{-1}	σ	c_{kGR}, gAPI	σ	c_{kLLd}, ohm m	σ	$c_{k,LLd/LLs}$ Centroids	σ
Data Set 1: San Vito 1								
1	111	9	207	14	1.8	0.8	0.80	0.08
2	73	5	248	24	97.0	113	0.93	0.27
3	143	8	194	7	15.0	6	1.57	0.21
4	88	8	276	35	5.3	2.6	0.82	0.10
5	106	13	189	18	2.2	2.1	1.07	0.13
6	92	8	168	21	3.2	1.6	0.83	0.09
Data Set 2: San Vito 3								
1	147	16	152	27	9.6	5.9	1.36	0.38
2	110	21	172	66	2.1	1.0	1.16	0.16
3	69	7	235	92	19.6	17.6	1.42	0.25
4	79	20	586	140	20.9	39	1.63	0.42

[a]The centroids are given as vectors $c_{k,v}$, where v is DT, GR, LLd, and LLd/LLs and k is the class number; σ is the standard deviation. Although log (LLd) is used in the analysis, we show the resistivity LLd instead for easier use in the discussion.

5. Materials and Methods

[19] We used the program FUZME version 2.1 [*Minasny and McBratney*, 2000], developed for the fuzzy classification of soil, to classify the logging data acquired in the two boreholes drilled in the San Vito area. The data have been separated into two distinct data sets (Table 2) according to the correspondence of depth and measured properties. Data sets 1 and 2 correspond respectively to wells SV-1 and SV-3. The two wells have been separated in the classification because of the tectonic collapse that has affected the area following the Gauro eruption. Well SV-3 is drilled on the rim of the collapsing zone while well SV-1 is inside the caldera, and their respective stratigraphies are appreciably different [*Rosi and Sbrana*, 1987]. We then chose to process each well separately.

[20] For the fuzzy analysis to be carried out, certain parameters need to be chosen for each data set. These are the distance measure (or metric), the number of classes and the value of the fuzzy exponent ϕ.

5.1. Choice of Distance-Dependent Metric

[21] The choice of the metric will help in driving the performance of the fuzzy clustering to an optimum. The most frequently used metric is the Euclidian distance

$$d_{ik}^2 = \sum_{v=1}^{q} (x_{iv} - c_{kv})^2 = (x_i - c_k)^T (x_i - c_k), \tag{8}$$

which gives equal weight to all the variables. Therefore, if one variable has much larger variance than the others, the former will heavily determine the result. This is typically our case. For instance, the variance of the acoustic slowness (DT) is much lower than the one of the natural radioactivity (GR). As this effect is undesirable, it can be suppressed by prior standardization of the variables to zero mean and unit

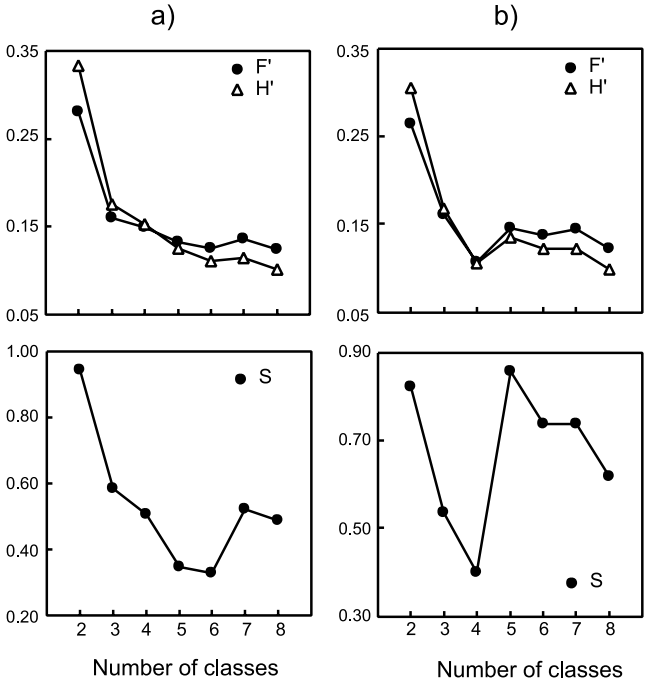


Figure 3. Determination of the optimal number of classes for each data set according to the minimum value of the fuzziness performance index (F') and of the modified partition entropy (H'). 6 classes were taken for data sets 1 and 4 classes for data set 2. The function S is also plotted against the number of classes and shows the same minimum for data sets 1 and 2.

variance, which is achieved through the use of the *Mahalanobis* [1930] norm, defined by

$$d_{ik}^2 = (x_i - c_k)^T \Sigma^{-1} (x_i - c_k), \tag{9}$$

where Σ is the sample variance-covariance matrix of the data matrix X. Using this metric, not only differences in variance but also correlation between variables are accounted for.

5.2. Degree of Fuzziness

[22] The degree of fuzziness is related to the internal degree of organization (or disorganization) of the data set. It determines the extent to which the final classes will be compact and separated. It is necessary to choose carefully the value of ϕ, so it keeps a balance between substructures in the data set and continuousness between the classes. Some workers use a value of 2 (*Bezdek* [1981]; *DeGruijter and McBratney*, [1988]). In our case, the classes are thought to have quite distinct properties, that should allow a relatively easy partitioning. However, each logging measurement averages a rather large sampling interval, this implying that each data has an intrinsic uncertainty that is not always within the variance of the class to which it belongs. A large ϕ will enhance this "data fuzziness" to the detriment of the interclasses fuzziness. On the basis of the work by *Odeh et al.* [1992] and *McBratney and de Gruijter* [1992], we tried different values of ϕ between 1.2 and 1.5. The best compromise was $\phi = 1.25$, as lower values gave a

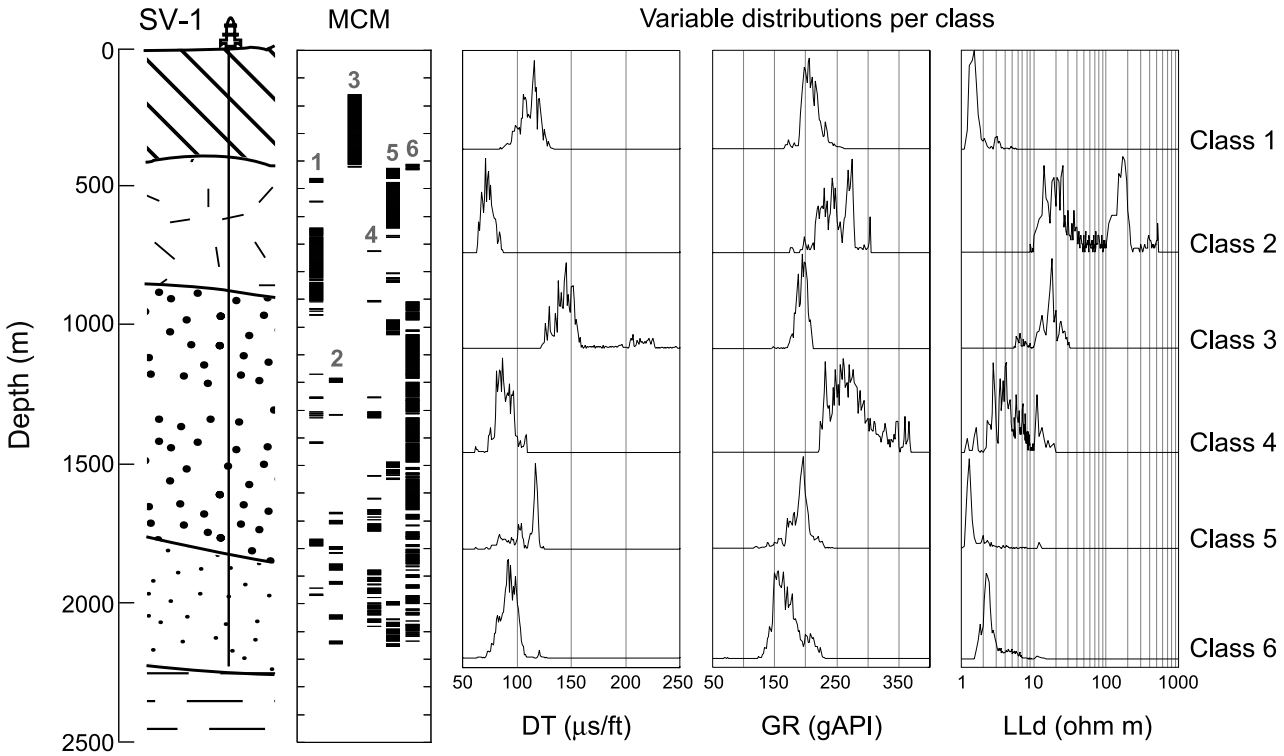


Figure 4. Results for well SV-1. From left to right, stratigraphy defined by *Rosi et al.* [1983], maximum class memberships (MCM) determined by the fuzzy classification, distribution of the physical properties: acoustic slowness (DT), natural gamma ray (GR), and electrical deep resistivity (LLd) inside the six classes.

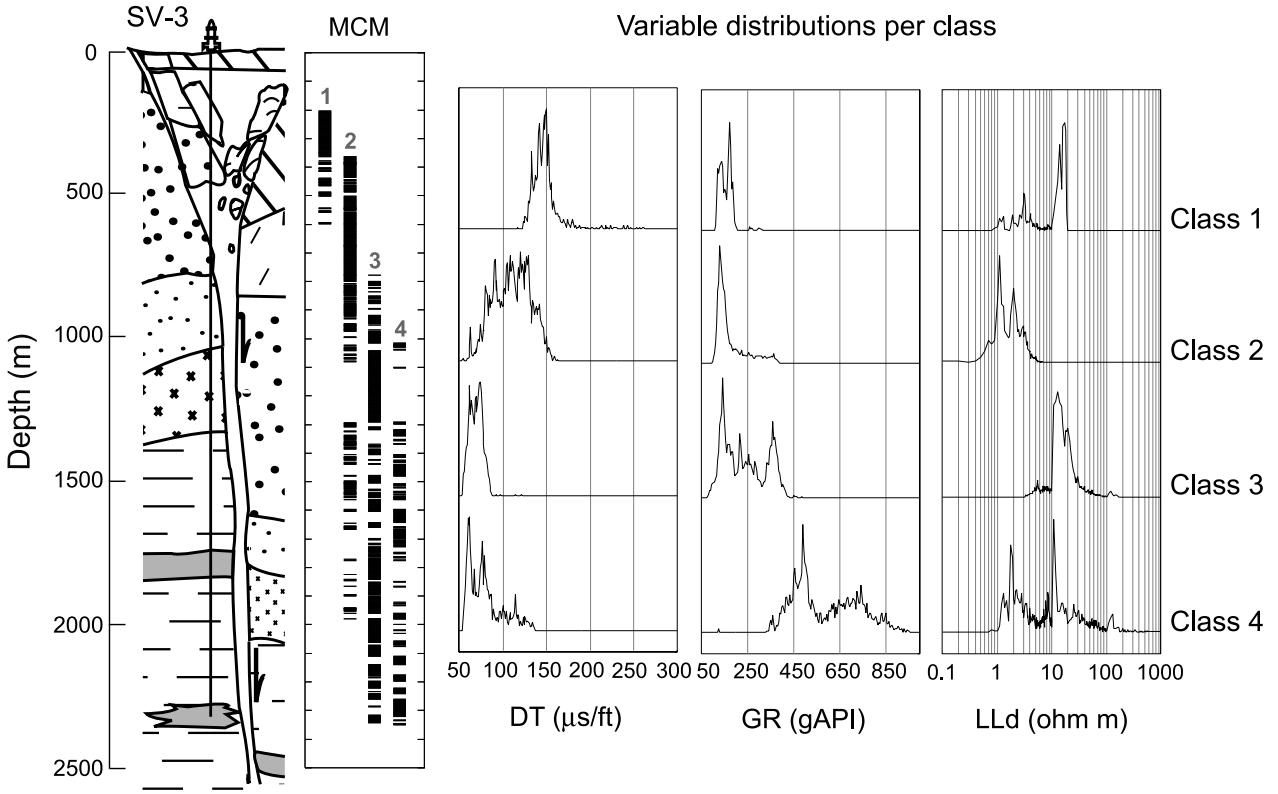


Figure 5. Results for well SV-3. From left to right, stratigraphy defined by *Rosi et al.* [1983], maximum class memberships (MCM) determined by the fuzzy classification, distribution of the physical properties: acoustic slowness (DT), natural gamma ray (GR), electrical deep resistivity (LLd) inside the six classes.

too hard classification, and higher values generated too much fuzziness between and inside the classes.

5.3. Optimal Number of Classes

[23] An optimal number of classes should reflect substructures in the data set. It is determined from prior knowledge of the data, and by using different mathematical functions devised to assess the validity of the solution. We use three of these functions in these study. The fuzziness performance index, noted F', estimates the degree of fuzziness generated by a specified number of classes. It is defined as [*Roubens*, 1982]

$$F' = \frac{1 - (p \times F - 1)}{F - 1}, \tag{10}$$

where F is the function [*Bezdek*, 1981]

$$F(M,p) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{p} \left(m_{ik}^{\phi}\right)^2. \tag{11}$$

As F' is a derivative of F (constrained as $0 \leq F' \leq 1$) as shown in equation (10), minimization of F' indicates an optimal number of fuzzy classes that best reflects the organization of the data set. As a consequence, $F' = 1$ corresponds to maximum fuzziness, and $F' = 0$ to nonfuzzy solution.

[24] The modified partition entropy, noted H', is defined as [*Roubens*, 1982]

$$H' = \frac{H}{\log p}, \tag{12}$$

where H is the entropy function

$$H = -\frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{p} m_{ik} \log(m_{ik}). \tag{13}$$

To validate the optimal number of fuzzy classes, one guesses that minimization of H (and hence H') equals to maximizing the amount of information about the substructures in X that is generated by the fuzzy k-means algorithm. Despite the heuristic nature of the use of these two validity functions, we will consider them as good indicators when the algorithm consistently identifies substructures within X for a given set of parameters $\{d, c_k, \phi\}$. For more definition and application of F' and H', one can refer to *McBratney and Moore* [1985] and *Odeh et al.* [1992].

[25] The last function has been devised by *Xie and Beni* [1991] and noted S. It measures the overall average compactness and separation of a fuzzy k partition. It is defined as

$$S = \frac{J_G(M,C)}{n^*(d_{\min})^2}, \tag{14}$$

where $J_G(M, C)$ is

$$J_G(M, C) = \sum_{i=1}^{n} \sum_{k=1}^{p} m_{ik}^{\phi} d_{ik}^2, \quad (15)$$

and $(d_{\min})^2$ is the separation measurement, where $d_{\min}$ is the minimum distance between cluster centroids, that is $d_{\min} = \min_{i,j}||c_i - c_j||$. In equation (14), the term $J_G(M, C)/n$ is called the compactness of the fuzzy partition of the data set. The function S is then defined as the ratio of the compactness to the separation of the fuzzy partition. A smaller S indicates a partition in which all the classes are overall more compact and better separate to each other.

[26] Using the same degree of fuzziness, $\phi = 1.25$, we performed the fuzzy k-means algorithm on the two data sets for $p = [2, 8]$. The choice of the range of p is based on initial expectations, i.e., the range should span an approximate maximum number of structurally plausible classes that can be gauged from the data or based on a priori knowledge. Plotting F', H', and S helps in determining the optimal number of classes (Figure 3). For data sets 1 and 2, S shows the best organization of the fuzzy k partition for 6 and 4 classes (S minimum), respectively, that are also relative minima for F' and H'. The plots for F' and H' decrease further when $p = 8$, but we shall discard this value. It brings little improvement of F' and H', while S is deteriorating, and a higher number of classes would complicate the interpretation of the classification. For data sets 1 and 2, all three functions clearly show a similar minimum at 6 and 4 classes, respectively.

6. Results

[27] The output of the fuzzy classification includes the membership for each data point in each class (given as probabilities). A class gathers all data points whose membership for this class is highest. A class k is characterized by its centroid, represented as a vector of coordinates (arithmetic mean) in the space of the variables. The scatter of a variable around the mean is indicated by the standard deviation σ. Another output of the fuzzy classification is the confusion index ξ_i, which measures for each data point the degree of class overlap in attribute space [*Burrough and McDonnell*, 1998]; ξ_i is calculated as $\xi_i = 1 - (p_{i,k_{\max}} - p_{i,k_{2\text{nd max}}})$, where $p_{i,k_{\max}}$ and $p_{i,k_{2\text{nd max}}}$ are the highest membership and the second highest membership, respectively, of data point i; ξ_i ranges from 0 (when the classes are well separated) to 1 (when they are not).

[28] Table 2 shows the properties of the class centroids and the corresponding standard deviations, for data sets 1 and 2. Figures 4 and 5attributed to well SV-1 and SV-3, respectively, show, from left to right: the lithology (determined from cuttings description [*Rosi et al.*, 1983]) the succession of the maximum class memberships (MCM), and the distributions of the data per class for the variables used in the classification. The meaning of a class will be commented by considering (1) that a centroid is representative of the properties of the class it refers to and (2) that the standard deviation indicates the degree of homogeneity of the class. From the succession of the maximum class memberships, we will define several "log units." A log unit may include one or more classes and separates sharp differences in the occurrence of the classes. The results for data sets 1 and 2 are presented separately in the following sections.

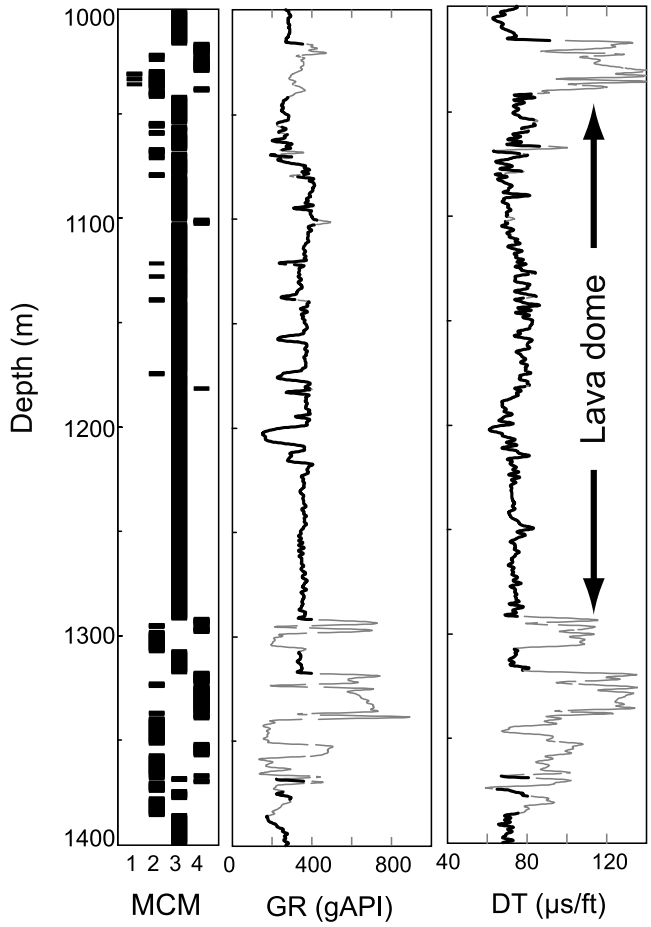


Figure 6. Thickness of the lava dome in SV-3. A constant low transit time (DT) and relatively high natural radioactivity (GR) between 1040 and 1290 m show the precise boundaries of the lava dome identified by *Rosi et al.* [1983].

6.1. San Vito 1

[29] In the first 260 m (between 160 and 420 m) studied of the SV-1 well (Figure 4), only class 3 appears. This class is characterized by the highest transit times (lowest acoustic velocities), showing a bimodal distribution. The first peak of the distribution, holding most of the data, is centered around 145 μs ft^{-1} (1 μs ft^{-1} = 3.28 μs m^{-1}), while the second is at 210 μs ft^{-1}. The natural radioactivity signature is very sharp around a relative average value of 200 gAPI. (The definition of the gAPI unit of radioactivity comes from an artificially radioactive formation, constructed at the University of Houston facility by the American Petroleum Institute, containing approximately 4% K, 24 ppm Th, and 12 ppm U. Its radioactivity was defined to be 200 gAPI units.) This class is one of the most resistive, showing a highly resistivity contrast (superior to one) between the deep and shallow resistivities (see Table 2), indicating that this formation is permeable. By comparing with the lithostratigraphy, this class can be associated to the pyroclastic incoherent deposits.

[30] The next geological unit, described as the Gauro's yellow tuffs, corresponds to two successive classes, class 5

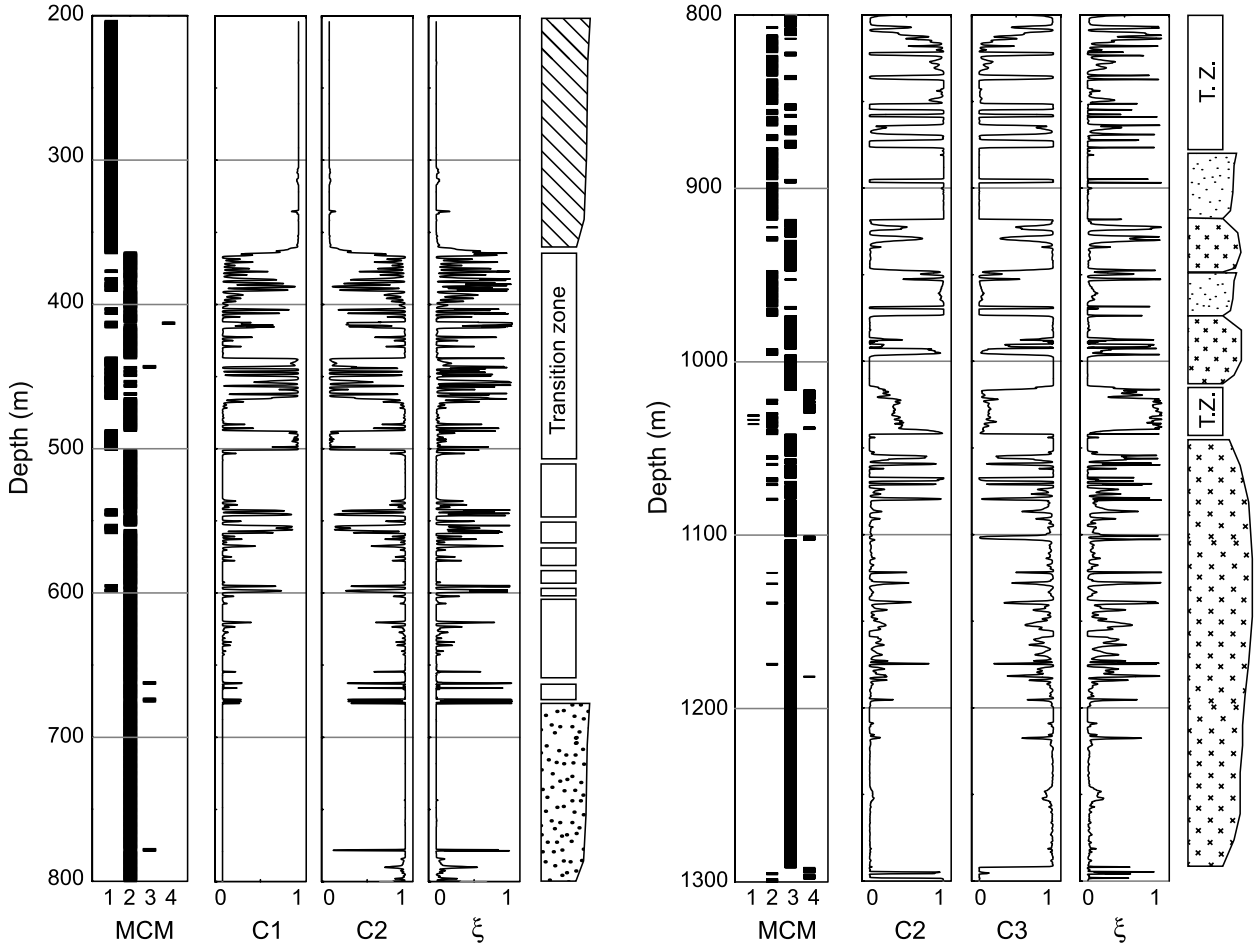


Figure 7. Information given by the confusion index (ξ) in SV-3. A high ξ indicates the transition zones (TZ), while it is low or zero in well-defined layers with distinct physical properties. C1, C2, and C3 indicate the class memberships for classes 1, 2, and 3, respectively. In the last column the patterns are the same as in Figure 1.

and then class 1. Their main difference is the resistivity contrast, superior to one in class 5, and inferior to one in class 1. They have similar deep resistivity and natural gamma ray distributions. Their deep resistivity and transit time are significantly lower than the previous unit. Class 5 and class 1 have respective thicknesses of 220 and 270 m.

[31] Class 6 is the main class between 910 and 1650 m and corresponds to the tuffites deposit. Its properties are well defined with moderate standard deviations. The transit time follows the decreasing trend observed from the top of the section, and is centered at 92 μs ft^{-1}. Its natural radioactivity distribution is shifted toward the lowest values of the entire hole (average of 168 gAPI). Its resistivity is higher than the one of the two previous classes, still it stays in the same order of magnitude.

[32] After 1650 m, we observe a mixture of classes 6, 4, 5, and 2. Class 2 appearance is rare and rather brief, showing the lowest and best defined transit time values (73 ± 5 μs ft^{-1}). Its bimodal deep resistivity signature, with one peak centered around 20 ohm m and another peak at 180 ohm m, can be partly explained by the vertical resolution limits of the resistivity tool (around 60 cm). Class 4 shows the same acoustic slowness as that of class 6, but has a higher deep resistivity and natural radioactivity, which by the way is the highest and most scattered of the section (276 ± 35 gAPI). The highest class diversity begins around 1800 m, corresponding to the beginning of the geological unit described as the chaotic tuffs, which indeed is composed of interlayered thin beds of different physical properties.

6.2. San Vito 3

[33] As *Rosi and Sbrana* [1987] did not clearly define the lithologies encountered during the first 800 m of SV-3, we will not relate, at this level, a class to a particular lithology. In Figure 5, class 1 is the first and only class to appears in SV-3 between the top and 365 m. It is then mixed with class 2 from 365 to 600 m. Class 1 shows the highest transit time values, with an average well-constrained around 147 μs ft^{-1}. The transition with class 2 is gradual and is mainly characterized by a decrease of the deep resistivity by 1 order of magnitude. Class 2 shows also a lower and broader transit time distribution. It is the sole class characterizing the interval between 600 and 800 m.

[34] The transition between class 2 and 3 is 250 m thick, and seems to correspond to the chaotic tuffs unit defined by

Rosi et al. [1983]. As in well SV-1, its average physical properties are difficult to determine.

[35] Class 3, with rather well-constrained low transit time and high resistivity values, clearly corresponds to the lava deposits [*Rosi et al.*, 1983]. The thickness of the lava dome can be precisely estimated at 250 m. Figure 6 illustrates how it is possible to determine with accuracy its position in the stratigraphy and its own physical properties. The lava dome is located between 1040 and 1290 m deep. It is a homogeneous formation, characterized by a low transit time around 70 μs ft^{-1}, an average deep resistivity of 17 ohm m and a natural radioactivity centered around 400 gAPI, which corresponds to the second peak of the natural radioactivity distribution.

[36] Following the presence of class 3 alone, the bottom part of SV-3 is composed of a succession of intercalated beds showing distinct physical properties, namely, that of classes 2, 3, and 4. Their major difference is the natural radioactivity, much higher in class 4. Class 2 is visible at the top of this log unit, but quickly becomes sparse. From the log curves, an increasing downward trend is visible in class 4 for the natural radioactivity and the deep resistivity. This unit will be documented with more detail in the discussion part. For this well, the ratio LLd/LLs is always greater than one, and, therefore, it was not used as a discriminant parameter.

7. Discussion

[37] Using the probabilistic character of this clustering, which allows to differentiate intercalations of well-defined layers from mixed lithologies characterizing disturbed transition zones, we discuss the physical meaning of each log unit and attempt to correlate the results found in each well of San Vito area.

7.1. San Vito 3

[38] We attempt here to precise the stratigraphy of the first 800 m of SV-3 well, which are not clearly defined in the vertical cross section represented in Figure 1. Making the correspondence of the classes found in the SV-1 and SV-3 wells, we notice that class 3 in SV-1 and class 1 in SV-3 have similar physical properties and correspond to the incoherent pyroclastic rocks. A 235-m-thick zone follows, composed mainly by a mixture of classes 1 and 2. A clue to characterize this zone is given by the confusion index ξ. Figure 7 shows the memberships for classes 1 and 2, for the depth interval 200–800 m, along with the confusion index. A high ξ characterizes the transition zone between 365 and 675 m, while it is mostly null in the two well-defined surrounding intervals. This transition zone corresponds probably to the mixed breccia described by *Rosi et al.* [1983] and filling the rim of the collapse zone.

[39] Class 6 of SV-1 has the closest physical properties of class 2 in SV-3. Its slightly higher *P* wave velocity and deep resistivity are probably due to its greater depth of burial. In SV-1, the interval where class 6 is prevalent (approximately between 900 and 1900 m, Figure 4), matches the chaotic tuffites deposits. *Rosi et al.* [1983] indicates that SV-3 crosses a rather thin layer of chaotic tuffites (Figure 5). These two observations agree to position the chaotic tuffites layer between 600 and 800 m in well SV-3.

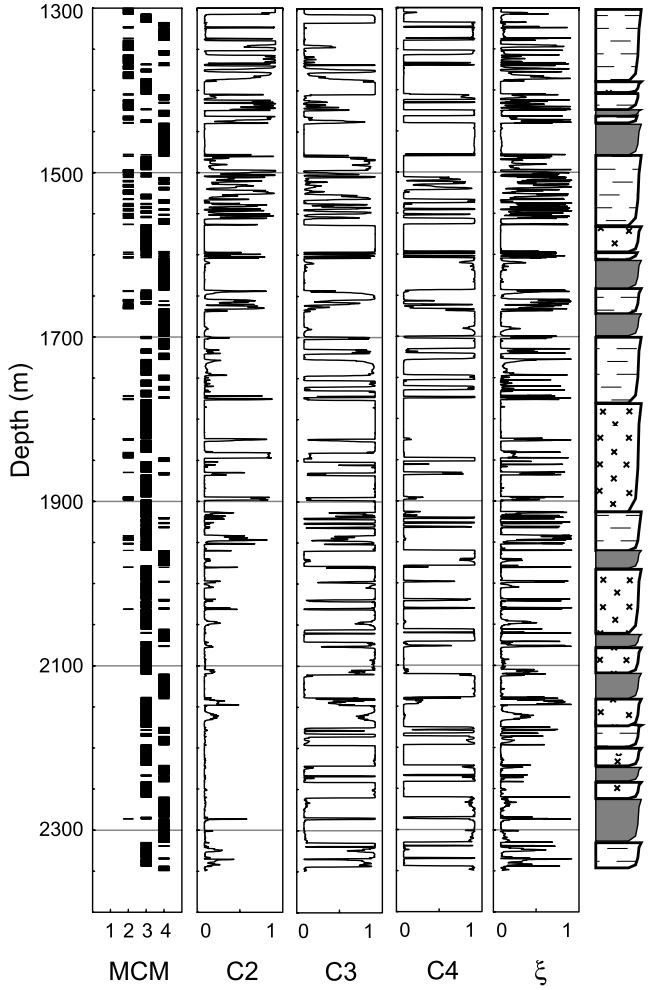


Figure 8. Information brought by the fuzzy classification: Maximum class memberships (MCM) and the confusion index (ξ) allows to describe precisely the stratigraphy of the thermometamorphic interval crossed by SV-3. We tentatively correlated class 3 to the tuffs-tuffites-lavas layers, and class 4 to the volcanosedimentary complex intervals. TZ, transition zone. In the last column the patterns are the same as in Figure 1.

[40] A second transition zone, defined by the simultaneous occurrence of classes 2 and 3, is found between 800 and 1040 m. The memberships of classes 2 and 3 and the confusion index ξ of this interval are represented in Figure 7 and can be divided into three parts. The first 80 m show a high confusion index and an undefined mixture of the two classes. This interval can be defined as a real transition zone. The following 120 m are made of well-characterized interbedded layers defined by class 2 or class 3. Inside each layer, ξ equals zero and the corresponding class membership is maximum, while its boundaries display a sharp increase of the confusion index. This interbedded layers zone corresponds well with the lower part of the chaotic tuffs deposit. The main differences between the layers are their transit time and deep resistivity (see Figures 4 and 5). The last 20 m are, as the previous transition zones, not representative of a clear formation, and mark the transition between the chaotic tuffs and the lava dome. From Figure 7, we see that in the first part of the dome, ξ varies between 0 and 0.9.

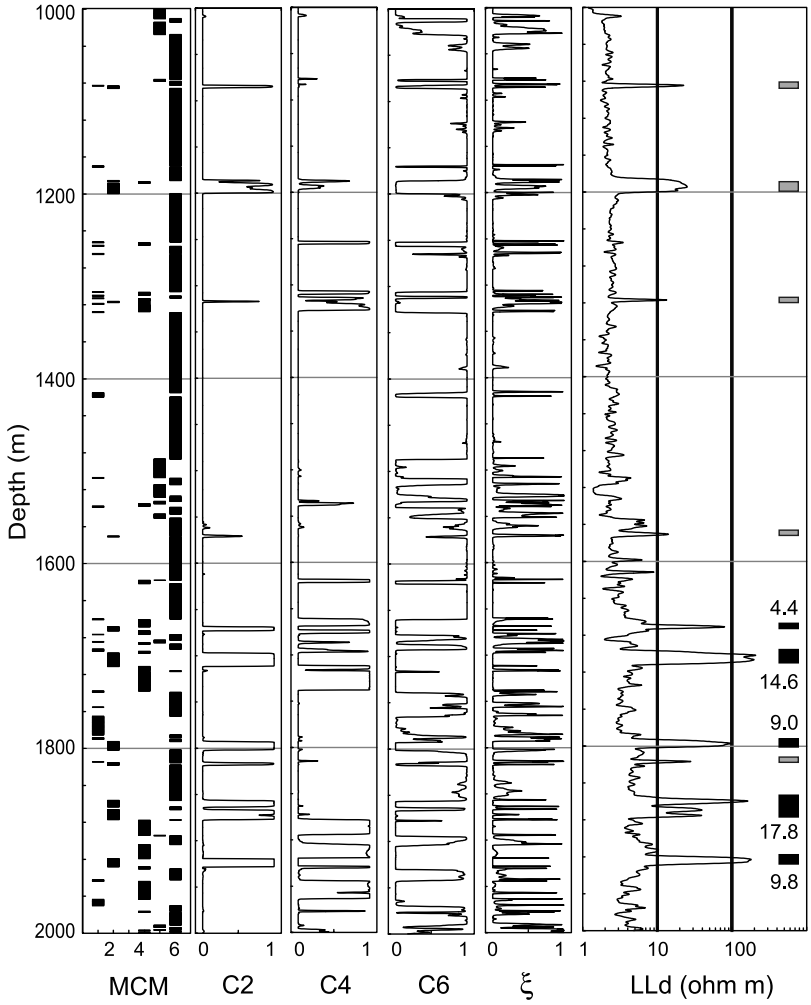


Figure 9. Information given by the confusion index (ξ) in SV-1. Left to right: Maximum class memberships (MCM), class memberships for classes 2 (C2), 4 (C4) and 6 (C6) and high readings of the deep resistivity (LLd) of class 2. A high ξ indicates the transition zones, while it is low or zero in well-defined layers with distinct physical properties. Typically, the high readings in deep resistivity that corresponds to a zero ξ match probably the lava beds described by *Rosi et al.* [1983]. We indicate their position (black rectangles) and estimated thickness.

These high confusion index zones may correspond to less homogeneous or more altered parts inside the lava body.

[41] From 1290 m to the bottom of the well, the last log unit is represented by a succession of classes 2, 3, and 4. Between 1290 and 1560 m, the confusion index is generally high and does not allow to clearly discriminate between deposits (Figure 8). After 1560 m, class 2 almost disappears, and the layers are better defined, showing an intercalation with an average thickness of 20 to 30 m. The maximum ξ peaks define the transition between the layers and allow to estimate the thickness of each layer. In the last column of Figure 8, we tentatively interpreted the succession of the class memberships. The pattern of the tuffs-tuffites-lavas layers was given when the confusion index was to high to differentiate each deposit. Although we attributed to class 3 the pattern of the lava, which has been previously clearly identified, this class can be, in this interval, interpreted either as lavas layers or as indurate layers of pyroclastic rocks. We correlated class 4 to the volcano-sedimentary complex intervals. The high natural radioactivity of this formation may be related to the severe hydrothermal alteration that has deeply transformed the nature of the formation [*Rosi and Sbrana*, 1987]. This thick unit, corresponding to the precaldera products, is only encountered at greater depth in well SV-1. This is certainly the reason why the very high values of natural gamma ray are never reached in the distributions obtained in SV-1.

7.2. San Vito 1

[42] SV1-discuss In well SV-1 we observe that each change in deposit is correlated with a different class. As a consequence, each class can be directly assigned to a volcanic deposit. The additional partitioning inside the yellow tuffs does not bear a significant change in terms of physical properties between classes 5 and 1. These classes show similar transit time, natural gamma ray and deep resistivity distributions and have only different LLd/LLs ratios. The boundary at 640 m marks actually the limit above which the LLd/LLs ratio is superior to 1 (classes 3 and 5), and below which it is inferior to 1 (every other

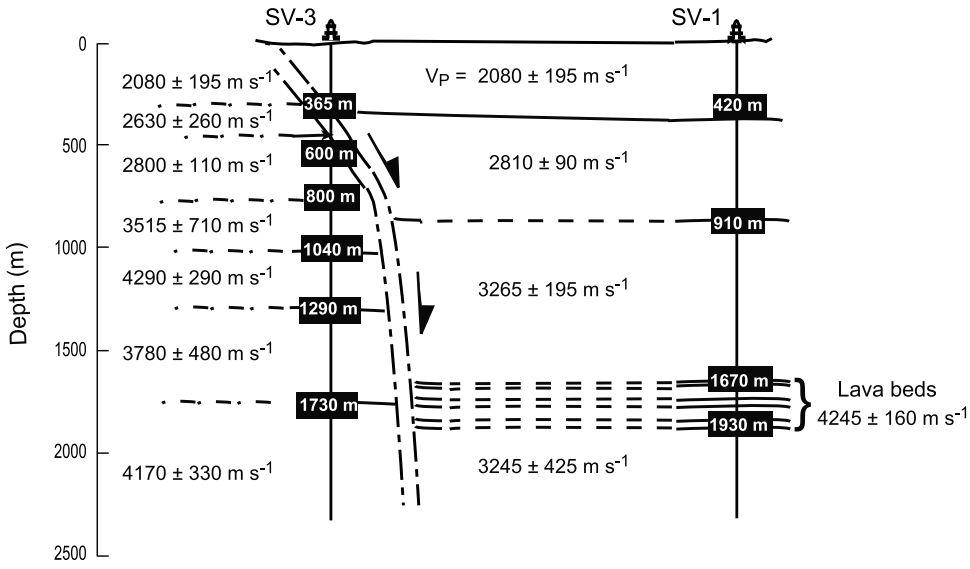


Figure 10. Two-dimensional compressional acoustic velocity structure of the San Vito area. The acoustic velocities are calculated as the inverse of the transit time value of the centroid of the class. In case of an interbedded interval (several classes describing the interval), we used a weighted average of the values of the centroids.

classes). The fact that the invaded and noninvaded zones have a similar resistivity (LLd/LLs = 1) is often interpreted by the impervious character of the geological unit. This is then probably the case of class 5.

[43] Below 910 m, the main class to appear is class 6 which is sometimes cut by thin and rare layers belonging to classes 1 and 2 and by thicker layers of classes 4 and 5. As observed in Figure 9, only the occurrences of classes 2 and 4 below 1600 m are associated with maximum class memberships, thus leading to a zero confusion index. The last column of Figure 9 shows the deep resistivity log curve between 1000 and 2000 m, where each occurrence of class 2 is correlated with a more or less broad peak of the resistivity. The resistivity around 20 ohm m belongs to the first four layers located at the top of the section. The resistivity around 180 ohm m corresponds to the true resistivity of class 2.

[44] In the deep resistivity log curve the occurrence of class 2 alone corresponds to a mean value of 180 ohm m. More than being the most resistive, class 2 has the highest compressive velocities, and a high natural radioactivity. These physical properties are typical of the lavas encountered in SV-3 (class 3). Class 2 may show lavas layers interbedding the precaldera chaotic tuffs and the postcaldera chaotic tuffites, as described by *Rosi et al.* [1983]. The thickness of these lava beds can be estimated with accuracy and are indicated in Figure 9. Class 4 occurrences are more difficult to associate to a particular lithology. However, because of a higher natural radioactivity, its class memberships and confusion index show that these occurrences are clearly distinct from class 6 (Figure 9). This high radioactivity may come from the presence of a particular secondary mineral holding a lot of potassium (like K-feldspar or illite [see *Rosi et al.*, 1983]).

7.3. The 2-D Velocity Structure

[45] Previous studies on the seismic velocity structure of the Phlegrean Fields are rare and the most significant are certainly the ones by *Aster and Meyer* [1988] and *Aster et al.* [1992]. They reconstructed a three-dimensional velocity structure in the central Campi Flegrei caldera, using a three-dimensional simultaneous inversion of *P* and *S* wave arrival times of 228 well-located microearthquakes recorded from February to June 1984. The dimension of the investigated area was 8 km along a east-west line, 6 km along a north-south line and 3 km in depth, centered 1 km north to the city of Pozzuoli.

[46] The study presented here differs from those of *Aster and Meyer* [1988] by several points: (1) the location of the investigated area; (2) the frequency used to explore the geological formations (10 kHz for downhole transit time measurements compared to 100 or 200 Hz for seismic data); and (3) the continuous character of the well-logging measurements (one measurement every half a foot), allowing to fully describe the investigated subsurface, and providing detailed information on the heterogeneity of the deposits, by opposition to the resolution of the seismic velocity model which depends on the ray path density.

[47] Figure 10 shows the 2-D acoustic velocity structure of the San Vito area, where the acoustic velocity representative of each unit and the depth of each interface are indicated. The *P* wave velocities are calculated by a weighting average of the transit time values of the centroids of the classes that are present in each interval. The weight is simply the relative amount of data belonging to each class in the interval.

[48] Because of the collapse between SV-1 and SV-3, only the first unit shows similar *P* wave velocities in the two wells. In the pyroclastic rocks, the velocity varies between 2080 m s^{-1} at the surface, where the rocks are incoherent, and 3400 m s^{-1} where they have suffered from hydrothermal alteration. The stronger values around 4440 m s^{-1}, belong to the lava body. SV-1 and SV-3 show significant difference in their *P* wave velocity profiles. In SV-1 the velocities increase regularly with depth from 2080 m s^{-1} at

surface to 3400 m s^{-1} around 2.5 km depth. In SV-3 the acoustic velocities model is more heterogeneous. This is partly due to its lithostratigraphy, composed by pyroclastic rocks and lavas.

8. Conclusion

[49] We applied a fuzzy classification technique on several well-logging data sets coming from two geothermal wells drilled in the San Vito area. This technique enabled partitioning the data sets into classes having the same physical properties and gave the average physical properties and the degree of homogeneity of each class. The additional concept of fuzziness with respect to the classical hard clustering, allowed us to take into account the continuous character of the subsurface geological and geophysical structures. Using this valuable information, we emphasized the geophysical structure of the volcanic deposits of the San Vito plain, which represents an area of recent collapse within the oldest Phlegrean caldera.

[50] The fuzzy classes were found to closely reflect the lithostratigraphy defined by *Rosi et al.* [1983] and described in detail by *Rosi and Sbrana* [1987]. The first and shallow deposit of incoherent pyroclastic rocks was recognized, independently, in SV-1 and SV-3, where it shows similar sharp distributions and average properties. It is one of the most resistive layers (10–15 ohm m), and it does not show any particular natural gamma ray signature. It has the lowest acoustic velocity of the whole section (centered around 2080 m s^{-1}).

[51] In SV-1 the thickness of the yellow tuffs has been precisely estimated at 490 m. Its deep resistivity distribution, mainly concentrated between 1 and 2 ohm m, allows to discriminate this unit from the rest of the formation. The acoustic velocities of the yellow tuffs are well defined, with values around 2810 ± 90 m s^{-1}.

[52] In SV-3 a transition between two geological units or collapse zones, represented by a chaotic mixture of deposits, could be differentiated from well-defined interbedded deposits. Particularly, SV-3 cuts, between 365 and 600 m, a mixed zone, corresponding probably to the rim of the collapse structure.

[53] Several lava beds belonging to the chaotic tuffs deposit in the SV-1 have been clearly recognized, among the chaotic tuffs and tuffites deposits. Their thicknesses have been estimated and vary from 4 to 18 m. The lava shows the highest *P* wave velocities associated with the smallest standard deviation (4245 ± 160 m s^{-1}). They have a high natural radioactivity around 250 gAPI. Their resistivity is 1 order of magnitude higher that the rest of the formation. In SV-3 the thickness of the lava dome was precisely estimated at 250 m.

[54] **Acknowledgments.** This is a CNRS-INSU-PNRN contribution (Thème risques volcaniques), and a EEC contribution of the Environmental and Climate Work programme (Volcanic risk). IPG contribution 1838.

References

Agip, Geologia e geofisica del sistema geotermico dei Campi flegrei, technical report, SERG-MESG, San Donato Milanese, Italy, 1987.

Aster, R. C., and R. P. Meyer, Three-dimensional velocity structure and hypocenter distribution in the Campi Flegrei caldera, Italy, *Tectonophysics*, *149*, 195–218, 1988.

Aster, R. C., R. P. Meyer, G. de Natale, A. Zollo, M. Martini, E. Del Pezzo, R. Scarpa, and G. Iannaccone, Seismic investigation of the Campi Flegrei: A summary and synthesis of results, in *Volcanic Seismology*, edited by P. Gasparini, R. Scarpa, and K. Aki, *IAVCEI Proc. Volcanol.*, *3*, 462–483, 1992.

Bezdek, J. C., Numerical taxonomy with fuzzy sets, *J. Math. Biol.*, *1*, 57–71, 1974.

Bezdek, J. C., *Pattern Recognition With Fuzzy Objective Function Algorithms*, Plenum, New York, 1981.

Bezdek, J. C., R. Ehrlich, and W. Full, FCM: The fuzzy *c*-means clustering algorithm, *Comput. Geosci.*, *10*(2–3), 191–203, 1984.

Burrough, P. A., and R. A. McDonnell, *Principles of Geographical Information Systems*, Oxford Univ. Press, New York, 1998.

DeGruijter, J. J., and A. B. McBratney, A modified fuzzy *k*-means method for predictive classification, in *Classification and Related Methods of Data Analysis*, edited by H. H. Bock, pp. 97–104, Elsevier Sci., New York, 1988.

Di Maio, R., D. Patella, Z. Petrillo, G. G. C. Siniscalchi, and P. Di Martino, Application of electric and electromagnetic methods to the definition of the Campi Flegrei caldera (Italy), *Ann. Geofis.*, *43*(2), 375–390, 2000.

Di Vito, M. A., R. Isaia, G. Orsi, J. Southon, S. de Vita, M. D'Antonio, L. Pappalardo, and M. Piochi, Volcanism and deformation since 12,000 years at the Campi Flegrei caldera (Italy), *J. Volcanol. Geotherm. Res.*, *91*, 221–246, 1999.

Dunn, J. C., A fuzzy relative of the isodata process and its use in detecting compact, well-separated clusters, *J. Cybernet.*, *3*, 22–57, 1974.

Ellis, D. V., *Well Logging for Earth Scientists*, Elsevier Sci., New York, 1987.

Forgy, E., Cluster analysis of multivariate data: Efficiency versus interpretability of classifications, *Biometrics*, *21*, 368 pp., 1965.

Hartigan, J. A., *Clustering Algorithms*, John Wiley, New York, 1975.

Harvey, P. K., and M. A. Lovell, *Core-Log Integration*, *Spec. Publ. Geol. Soc.*, *136*, 422 pp., 1998.

Hearst, J. R., and P. N. Nelson, *Well Logging for Physical Properties*, 571 pp., McGraw-Hill, New York, 1985.

Hurst, A., M. A. Lovell, and A. C. Morton (Eds.), *Geological Applications of Wireline Logs*, vol. 1, *Spec. Publ. Geol. Soc.*, *48*, 1990.

Hurst, A., C. M. Griffiths, and P. F. Worthington (Eds.), *Geological Applications of Wireline Logs*, vol. 2, *Spec. Publ. Geol. Soc.*, *65*, 1992.

Lofts, J. C., Integrated geochemical-geophysical studies of sedimentary reservoir rocks, Ph.D. thesis, Univ. of Leicester, Leicester, U.K., 1993.

Mahalanobis, P. C., On tests and measures of groups divergence, I, *J. Asiatic Soc. Benagal*, *26*, 541, 1930.

McBratney, A. B., and J. J. de Gruijter, A continuum approach to soil classification by modified fuzzy k-means with extragrades, *J. Soil Sci.*, *43*, 159–175, 1992.

McBratney, A. B., and A. W. Moore, Application of fuzzy sets to climatic classification, *Agric. For. Meteorol.*, *35*, 165–185, 1985.

McQueen, J. B., Some methods for classification and analysis of multivariate observations, in *5th Berkeley Symposium on Mathematics, Statistics, and Probability*, vol. 1, edited by L. M. Le Cam and J. Neyman, pp. 281–298, Univ. of Calif. Press, Berkeley, 1967.

Minasny, B., and A. B. McBratney, FuzME version 2.1, Aust. Cent. for Precision Agric., Univ. of Sydney, New South Wales, 2000. (Available at http://www.usyd.edu.au/su/agric/acpa).

Odeh, I. O. A., A. B. McBratney, and D. J. Chittleborough, Soil pattern recognition with fuzzy-*c*-means: Application to classification and soil-landform interrelationships, *Soil Sci. Soc. Am. J.*, *56*, 505–516, 1992.

Orsi, G., S. de Vita, and M. A. di Vito, The restless, resurgent Campi Flegrei nested caldera (Italy): Constraints on its evolution and configuration, *J. Volcanol. Geotherm. Res.*, *74*, 179–214, 1996.

Rabaute, A., L. Briqueu, P. K. Harvey, and H. Mercadier, Characterisation of oceanic sediments and basalts from Hole 833B, New Hebrides island arc (Vanuatu) from downhole measurements, *Sci. Drill.*, *6*(2), 91–101, 1997.

Rittmann, A., L. Vighi, F. Falini, V. Ventriglia, and P. Nicotera, Rilievo geologico dei Campi Flegrei, *Boll. Soc. Geol. Ital.*, *69*, 117–362, 1950.

Rosi, M., and A. Sbrana, Phlegraean Fields, *Quad. Ric. Sci.*, *114*(9), 175 pp., 1987.

Rosi, M., A. Sbrana, and C. Principe, The Phlegraean Fields: Structural evolution, volcanic history and eruptive mechanism, *J. Volcanol. Geotherm. Res.*, *17*, 273–288, 1983.

Roubens, M., Fuzzy clustering algorithms and their cluster validity, *Eur. J. Oper. Res.*, *10*, 294–301, 1982.

Ruspini, E. H., A new approach to clustering, *Inf. Clustering*, *15*, 22–32, 1969.

Ruspini, E. H., Numerical methods for fuzzy clustering, *Inf. Sci.*, *2*, 319–350, 1970.

Serra, O., *Fundamentals of Well-Log Interpretation*, Elsevier Sci., New York, 1984.

Xie, X. L., and G. Beni, A validity measure for fuzzy clustering, *IEEE Trans. Pattern Anal. Mach. Intel.*, *13*(8), 841–847, 1991.

Yven, B., Mineralogy, microstructure and physical properties of volcanic rocks from the Phlegrean Fields complex (Italy): Contribution of laboratory and downhole measurements to the description of its internal structure, Ph.D. thesis, 340 pp.,Univ. Paris, France, 2001.

Zadeh, L. A., Fuzzy sets, *Inf. Control*, *8*, 338–353, 1965.

Zamora, M., G. Sartoris, and W. Chelini, Laboratory measurements of ultrasonic wave velocities in rocks from the Campi Flegrei volcanic system and their relation to other field data, *J. Geophys. Res.*, *99*, 13,553–13,561, 1994.

W. Chelini, Agip S.p.A, Formation Evaluation Dpt., San Donato Milanese, Italy. (walter.chelini@agip.it)

A. Rabaute, Laboratoire de Géologie, Ecole Normale Supérieure, Rue Lhomond, F-75005 Paris, France. (rabaute@geologie.ens.fr)

B. Yven, Service Direction Scientifique/Geologie Geoprospective ANDRA, Chatenay-Malabry, France. (Beatrice.Yven@andra.fr)

M. Zamora, Département des Géomatériaux, IPGP, 4 Place Jussieu, F-75252 Paris cedex 05, France. (zamora@ipgp.jussieu.fr)