

## Article

# Geophysical Subsoil Characterization and Modeling Using Cluster Analysis for Seismic Microzonation Purposes

Patrizia Capizzi  and Raffaele Martorana \* 

Dipartimento di Scienze della Terra e del Mare, Università degli Studi di Palermo, 90123 Palermo, Italy; patrizia.capizzi@unipa.it

\* Correspondence: raffaele.martorana@unipa.it

**Abstract:** In the municipality of Enna, 80 HVSR measurements were performed, and some of these were combined with MASW seismic measurements, which made it possible to constrain the data inversion and obtain significant shear wave velocity models. A reconstruction of the depth of the seismic bedrock was performed for the whole territory, showing different depths for the higher and lower areas, as evidenced also by the Vseq parameter map. The frequency peaks identified in the H/V curve were analyzed through a cluster analysis algorithm to evaluate similarities that allow these peaks to be divided according to their stratigraphic origin. A non-hierarchical analysis algorithm modified in such a way as to avoid any a priori choice that could influence the partition has been used. The cluster analysis made it possible to divide the frequency peaks into five groupings, each of which was then associated with a seismic discontinuity, according to the geological contacts expected in the subsoil. Finally, the inversion of the data made it possible to reconstruct the geometries of these geological contact surfaces and to reconstruct a 3D model of the subsoil, which agrees well with the surface geology of the area.

**Keywords:** HVSR; MASW; cluster analysis; seismic microzonation



**Citation:** Capizzi, P.; Martorana, R. Geophysical Subsoil Characterization and Modeling Using Cluster Analysis for Seismic Microzonation Purposes. *Geosciences* **2023**, *13*, 246. <https://doi.org/10.3390/geosciences13080246>

Academic Editors: Claudia Pirrotta, Sebastiano Imposa, Maria Serafina Barbano and Sabrina Grassi

Received: 6 July 2023

Revised: 11 August 2023

Accepted: 12 August 2023

Published: 15 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, the horizontal-to-vertical spectral ratio (HVSR) technique applied to microtremors has acquired increasing importance for the evaluation of the seismic characteristics of a site because the HVSR can aid in identifying areas with similar seismic responses. Theoretical considerations [1] and experimental tests showed that the amplification of horizontal motions between the bottom and the top of a sedimentary cover is related to the ratio between the spectra of the horizontal and vertical components of the ground velocity [1,2]. This ratio is a measure of the ellipticity of Rayleigh wave polarization, overlooking the contribution of Love and body waves. Assuming that subsoil can be represented as a stack of homogeneous horizontal layers and imposing some geometric and/or physical constraints, it is possible to estimate the parameters of the shear wave velocity model [3,4]. Furthermore, the HVSR made it possible, in many cases, to reconstruct the top of the seismic bedrock [5,6]. The integration of data related to the HVSR and active techniques based on the analysis of surface waves can greatly reduce the uncertainties in the interpretation models.

In evaluating the resonance frequencies of the ground, a peak in the HVSR curve is not necessarily related to stratigraphic sources and therefore to a resonant frequency of buried structures. It could depend on the characteristics of the noise sources or complex phenomena of interference and focusing of P and S waves. Cluster analysis constitutes an important contribution to the attempt to identify and group peaks of HVSR curves relating to the same stratigraphic sources, topographical, anthropic or of any other nature. The hypothesis on the knowledge of the sources of the HVSR peaks is fundamental information for data inversion and correct modeling of the subsoil.

Cluster analysis [7] is a statistical procedure that can be used in the exploratory phase of a survey or when there are no a priori hypotheses. This analysis is widely used in different research fields [8,9], being a good tool when there is the need to classify a large amount of information into meaningful and manageable groups. In fact, it allows us to identify within a set of objects of any nature some subsets, called “clusters”, which are mutually exclusive and similar. The dataset is divided into groups according to the level of similarity of data, evaluated starting from the values of a series of pre-selected variables for each cluster. Before proceeding with the analysis, it is very important to know the strengths and weaknesses of the different clustering techniques, hierarchical and non-hierarchical, and to explore the reasons for their analytical differences. Generally, methods that are unconstrained by small variations in the data are preferred, particularly if the clusters are thought to be rather homogeneous. In this case, non-hierarchical analysis techniques are preferred.

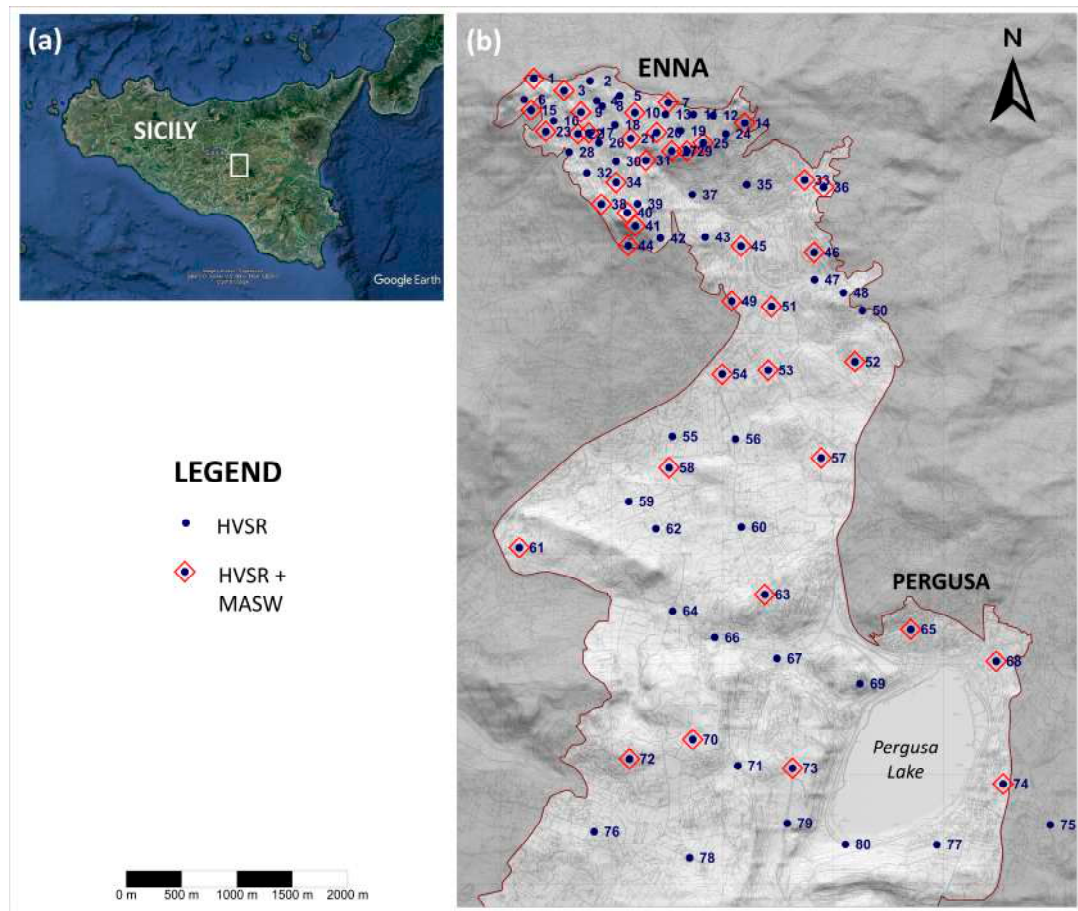
Cluster analysis can therefore prove to be very useful both in the statistical processing of the HVSR curves in order to extract stratigraphically interpretable curves [10] and in the interpretation of stratigraphic peaks with one-dimensional modeling.

Because the inversion of HVSR curves implies monodimensional distribution of  $V_s$ , before the inversion, the cluster analysis can be applied to the spatial distribution of all the HVSR stratigraphic peaks to group those probably due to the same stratigraphic interface [11–13] to define the seismic layers, each characterized by a specific shear wave velocity range. The data of each cluster can be interpreted by imposing conditions of maximum similarity between the 1D models relating to each measurement point. Furthermore, the uncertainty of the solution of the inverse problem in the HVSR can be reduced by using data from other geophysical methods as a priori constraints [14–20]. In particular, the surveys performed with the Multichannel Analysis of Surface Waves (MASW) method can be utilized to obtain one-dimensional shear wave velocity models that can be used as initial models for the inversion of the HVSR curves [11,21–24].

The depths of the seismic layers obtained by inverting the HVSR curves, when referring to the same cluster, can then be interpolated to reconstruct the three-dimensional trend of the seismic interface.

This work presents an application of cluster analysis on a dataset of 80 HVSR curves, aimed at the reconstruction of stratigraphic surfaces in the subsoil of Enna (Italy) (Figure 1). Microtremor data processing had previously been performed to extract stratigraphically interpretable curves using a clustering method proposed by [10]. From then, a dataset of HVSR peaks was obtained and analyzed with a non-hierarchical cluster analysis algorithm [6] to subdivide the peaks into clusters, each referring to a specific stratigraphic surface.

The data were inverted by constraining them with 40 MASW models and imposing conditions of maximum similarity for the HVSR inverse models belonging to the same cluster to subsequently be able to reconstruct the three-dimensional trend of the stratigraphic surfaces.

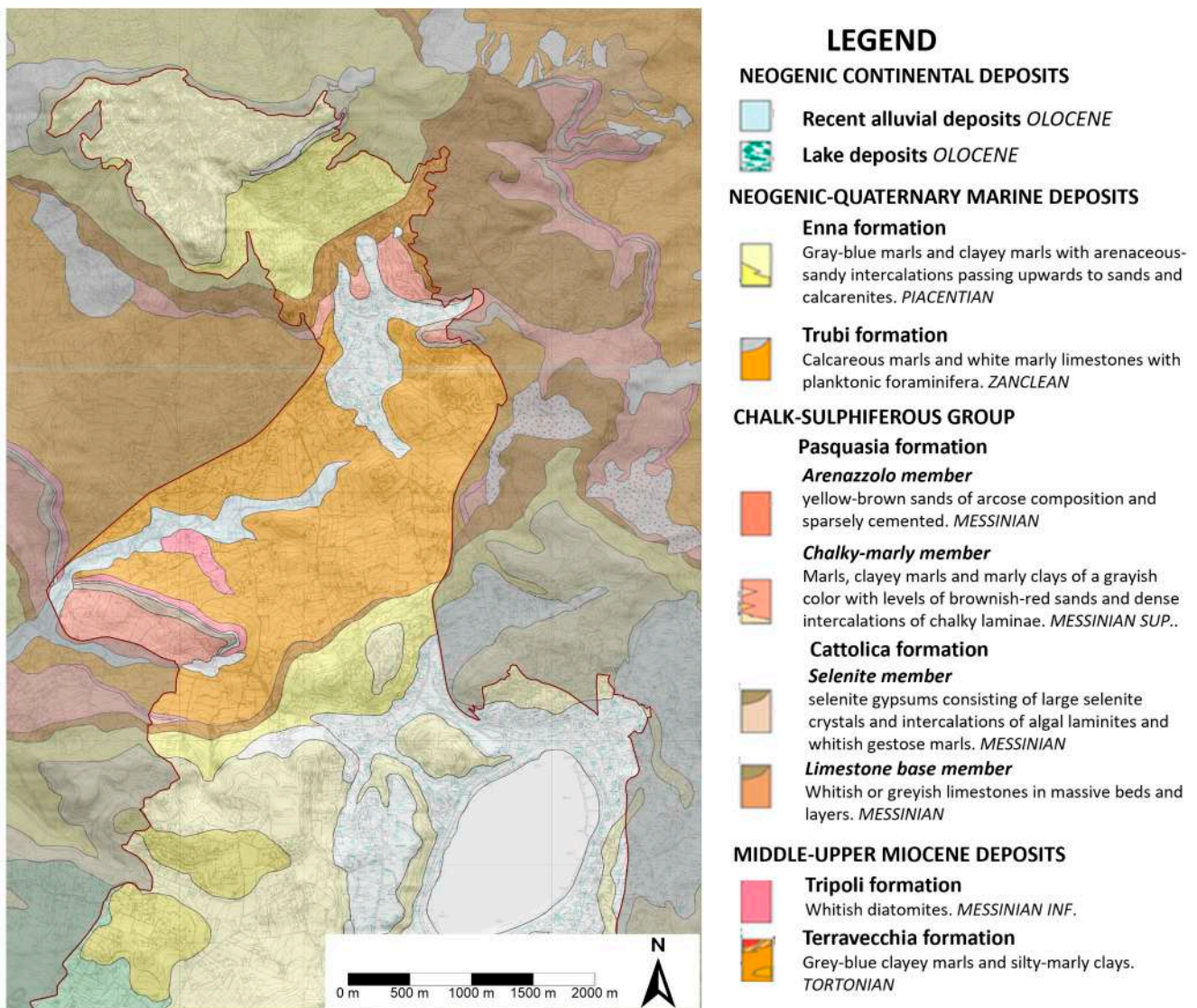


**Figure 1.** (a) Investigated area; (b) location of the measurement points carried out in the urbanized area of Enna. The brown line delimits the urbanized area, subject to seismic microzonation. The blue point indicates the positions of the HVSR surveys, and the red diamonds indicate the positions of the MASW surveys.

## 2. Geological Setting of the Studied Area

The study area is located in the central-southern sector of Sicily and is part of the Caltanissetta Basin, a vast structural depression characterized by large syn-tectonic sedimentation from the Upper Tortonian to the Plio-Quaternary. This basin was the ancient foredeep basin of the Apennine-Maghrebi chain (Late Miocene) and is currently part of the chain since it was incorporated into the accretionary prism [25].

In particular, the formations outcropping in the study area (Figure 2), starting from the bottom upwards, are made up of clayey marls of the Terravecchia Formation (Lower Tortonian–Lower Messinian) with maximum thicknesses of about 300–400 m. The gypsum-sulphiferous group follows (Lower Messinian–Upper Messinian) made up of diatomites of the Tripoli Formation (maximum thickness 30 m), the base limestones (about 50 m), the selenitic gypsums (about 50 m), the Cattolica Formation composed of rock salt with intercalation of kaolinite (about 500 m) and the Pasquasia formation (marls and clayey marls less than 100 m thick).



**Figure 2.** Geological map of the area of the municipality of Enna. The brown line delimits the urbanized area, subject to seismic microzonation.

Following upwards is the Trubi, consisting of an alternation of calcareous marls and white marly limestones (maximum thickness of about 100 m). The Enna Formation is found here, characterized by marl and clayey marl (ENNa lower member), which gradually pass upwards to the sands and calcarenites of Capodarso (ENNb upper member). The thickness is about 250 m (Piacenziano).

Finally, there are lacustrine and alluvial Quaternary deposits.

### 3. Materials and Methods

#### 3.1. HVSR Method

The horizontal-to-vertical spectral ratio (HVSR) technique [1,2], applied to ambient seismic noise, is a method widely used to estimate the resonance frequencies of geological structures and is today one of the most applied methods for expedite microzonation studies of large urban areas [26,27].

A large number of surveys based on this technique were carried out to study site effects in sedimentary basins, near faults or cavities or to estimate the resonance frequency of buildings [28–30].

The reliability of the HVSR technique shows some criticalities linked to some assumptions, generally not easily verifiable experimentally, related to the elastic properties of the subsoil, the composition of the microtremor in terms of superposition of surface or body waves and the space–time distribution of the sources.

For example, if the shape of the HVSR curves is mainly controlled by the S-wave transfer function of the shallowest sedimentary layers [1,2,4,31], then both the frequencies and amplitudes of the HVSR curve relative maxima are directly related to subsoil resonance frequencies and site amplification factors. On the other hand, if the shape of the HVSR curves is controlled by the ellipticity of the fundamental mode of Rayleigh wave polarization [32–34], then only an indirect correlation between HVSR curve peaks and site amplification may be supposed.

To obtain reliable and geologically interpretable HVSR curves, it is necessary to adopt robust criteria for a correct seismic noise recording and for the choice and analysis of an optimal set of analysis time windows [35]. There are several hypotheses on the choice of these criteria.

Strong transient noise generally depends on nearby sources and therefore should be removed, keeping only the low amplitude part of the recording, for a correct estimation of resonant frequencies [35,36].

The windows selection can be made in both time and frequency domain. In the time domain, windows are selected by visual inspection, while in the frequency domain the spectral ratio is calculated for each time window, and subsequently those HVSR curves which do not deviate too much from the average are selected.

In order to invert the HVSR curves, it is assumed that the subsoil has viscoelastic properties and is horizontally stratified. Furthermore, environmental noise is considered to be composed of the superposition of Rayleigh and Love waves generated by a random distribution of independent sources [3,16,37,38]. The inversion of the HVSR curves can be utilized to obtain 1D layered seismic velocity models. However, to limit the uncertainty, it must be constrained with a priori data achieved from active seismic methods, such as the MASW method [18,23,39,40]. The MASW data, if acquired from multi-station records, can be used to derive the ellipticity of the fundamental mode of Rayleigh waves [41,42] even if the typical frequencies of the sources used, generally higher than those of the microtremor, allow depths of minor investigation.

### 3.2. MASW Method

The Multichannel Spectral Analysis of Surface Waves (MASW) method [43] is an active investigation technique, based on the analysis of the geometric dispersion of surface waves, that allows the shear wave velocity in the subsoil to be estimated. Surface wave dispersion occurs when the elastic wave velocity varies with depth and, consequently, the spectral components of a surface wave assume different velocities depending on frequency. In the MASW method, an artificial seismic source, usually a sledgehammer, generates a wave front, which is recorded by a linear array of equally spaced geophones placed on the surface. The seismic traces are analyzed in the frequency and wave number domain, obtaining a spectral graph of the oscillation amplitude as a function of phase velocity and frequency. From this, depending on the orientation of the source and the geophones, it is possible to obtain the dispersion curves of the fundamental and higher modes of the Rayleigh and Love waves. The dispersion curves are inverted to achieve stratified one-dimensional models of the shear wave velocities down to investigation depths, which, depending on the analyzed frequency range, generally do not exceed 30 m.

The shear wave velocity models obtained by inverting the MASW data can be used to constrain the topmost part of the inverted HVSR models to ensure their interpretive robustness [23,40].

### 3.3. The Cluster Analysis

The cluster analysis is the procedure that can identify within a set of objects some subsets, called clusters, that tend to be homogeneous according to some criteria. The statistical units are divided into a number of groups according to their level of similarity (internal cohesion), evaluated from the values that a number of variables chosen takes in each unit. Generally, in the cluster analysis, it is not necessary to have in mind any interpretative model [44]. The partition is successful if the objects within the clusters are closer to each other than other in different clusters [45].

Many clustering algorithms exist [46,47] and can be categorized into two main types: hierarchical clustering (hc) and non-hierarchical clustering (NHC). The HC methods are explorative and are not necessary for defining a priori the number of clusters. The HC methods work with a measure of proximity between the objects to be grouped together. A type of proximity can be chosen which is suited to the subject studied and the nature of the data. One of the results of HC methods is the dendrogram, which shows the progressive grouping of the data. It is then easy to gain an idea of a suitable number of classes into which the data can be grouped.

NHC methods do not have a tree structure and generate new clusters by merging or dividing previous clusters, grouping data in order to maximize or minimize some evaluation criteria. The elements that most influence the results of the cluster analysis are the shape, size and number of clusters; the presence of outliers; the level of overlap between clusters and the type of similarity/dissimilarity measure chosen.

To evaluate the differences between the various clustering techniques, which can also produce results significantly different from each other, the best way is to assess how the different techniques reproduce the structure of known data. These assessments are typically performed on simulated data and are often difficult to interpret and may be contradictory.

Various studies [48,49] suggest that different grouping strategies often lead to not dissimilar results, while others highlight specific cases of strong divergence [44,47]. However, the criteria for choosing between the two types of algorithms (hierarchical and non-hierarchical) have not yet been sufficiently explored, and very different positions can be found in the literature.

#### Cluster Analysis of HVSR Data

Windows selection for the average HVSR curve estimation is generally performed through visual inspection of the HVSR curves as a function of time. Often it is very difficult to identify the correct time window to be used for the calculation of the mean HVSR. The lack of a not arbitrary selection criteria makes the result clearly operator-dependent and therefore not optimal. To overcome this problem, an agglomerative hierarchical clustering (AHC) algorithm can be used to extract frequency and amplitude of HVSR curves determined in sliding time windows [10].

After identifying the HVSR peaks attributable to site effects, the identification of areas with similar behavior in a seismic perspective involves the simultaneous evaluation of several parameters such as frequency and amplitude of peaks, positions and mutual distances between the measuring points. The task becomes particularly difficult when different peaks on single HVSR curves are identified.

In this regard, we used a procedure proposed by [6], which is based on a centroid-based NHC algorithm. NHC methods are limited by the need to choose a priori the number of clusters. However, the method used here does not fix the number of  $k$  clusters, but it chooses automatically for each  $k$  value tested the initial centroids from dataset. The parameters considered for the determination of the clusters are the coordinates, the peak frequency and amplitude, and the outcropping lithology at the measurement point. This latter is converted into numerical information by assigning the average density of the corresponding lithology, directly related to the seismic behavior.

The Euclidean normalized distance of each unit from the initial cluster centroid and those obtained after each iteration is calculated as the weighted sum of the distances of all the variables considered.

The choice of the optimal number of clusters and weights is made by testing different values of the number of clusters  $k$  and by analyzing for each the trend of the deviance within each cluster and between clusters [6,13].

### 3.4. Application to the Enna Territory

From October 2014 to February 2015, we performed a passive seismic campaign to determine the resonance frequency of the subsoil, as part of the seismic microzoning studies for the Municipality of Enna (Italy). In this regard, 80 measurement stations distributed throughout the Enna area were chosen (Figure 1). An HVSr analysis was carried out for each measuring station and MASW surveys were carried out at half of these points, aimed at the seismic characterization of the soil and at constraining the inversions of the HVSr curves.

#### 3.4.1. Dataset Acquisition and Elaboration

MASW surveys were performed at half (40) of the 80 HVSr measurement stations. The MASW data were used to obtain an estimate of the shear wave velocities in the surface layers, necessary for a constrained inversion of the HVSr data. The seismic traces had a sampling interval of 256  $\mu$ s and a recording time of 1024 ms.

For each survey, 24 vertical 4.5 Hz geophones were placed with spacing of 2 m, to investigate the vertical component of Rayleigh waves (ZVF). The source was a 5 kg sledgehammer. Shots were generated at offsets of 2, 4 and 7 m for both sides of the array. Of these shot records, the ones which showed a better signal-to-noise ratio were used for the inversion and interpretation. A stacking process on five shots was set out on all the recordings to increase the signal/noise ratio.

The data were processed and inverted using the winMASW software to obtain from the dispersion curves the vertical profile of the shear wave velocity  $V_s$ . Figure 3 shows an example of processing a MASW survey.

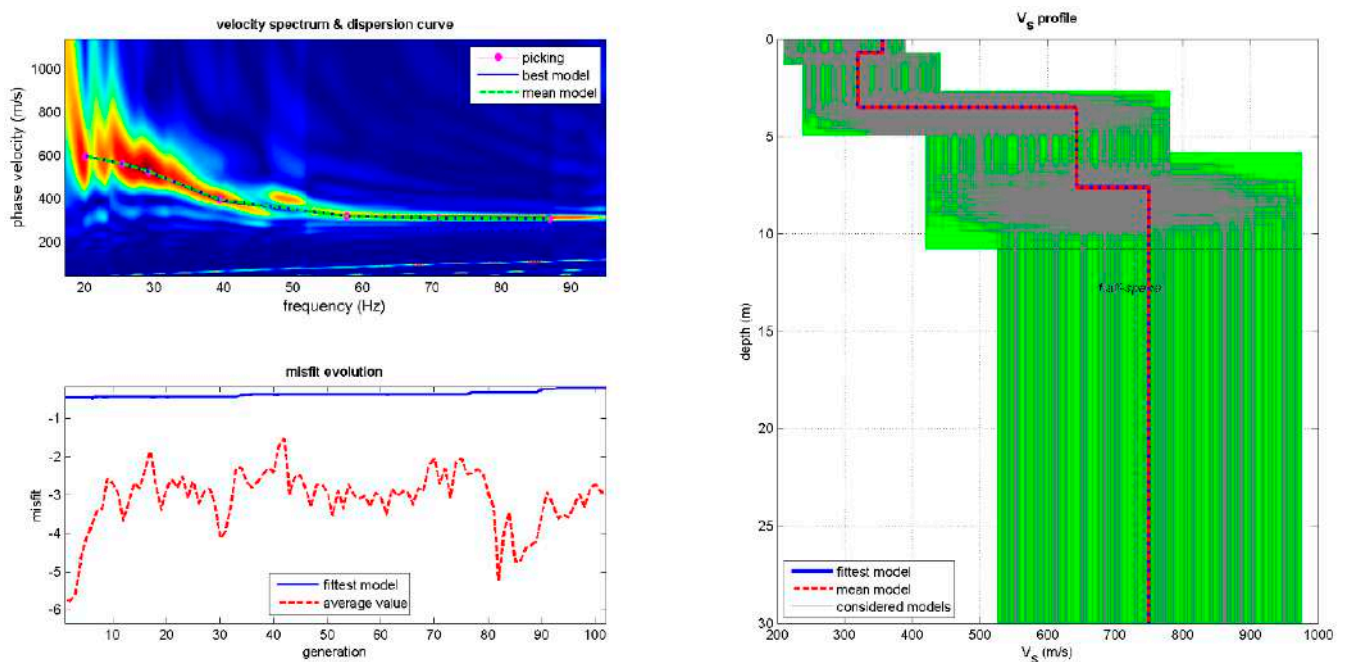
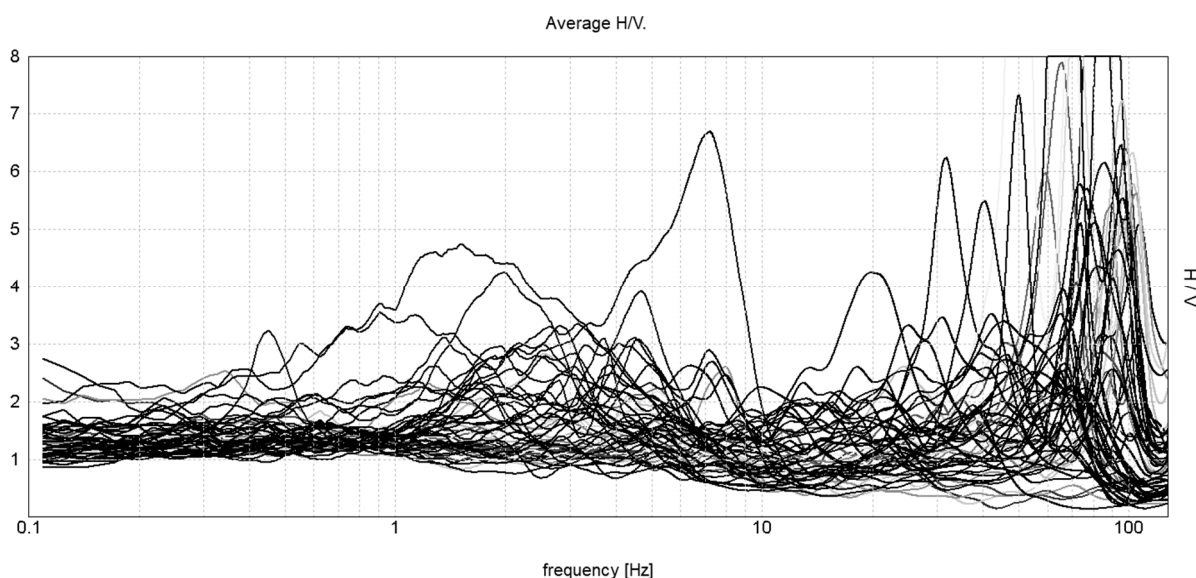


Figure 3. Example of elaboration of a MASW survey.

The HVSR measurements were performed using the TROMINO<sup>®</sup> (Venice, Italy) digital seismic detector from Micromed S.p.A. with a sampling frequency of 256 Hz and a recording duration of 46 min in order to obtain a useful signal of no less than 30 min [35].

The traces were processed using the Micromed S.p.A. Grilla software. For each recorded signal, the choice of the time windows to be analyzed, each lasting 50 s, was made manually, analyzing the graphs of the temporal and azimuth variation of the H/V spectrum and choosing only the noise time windows characterized by temporally stationary spectral estimates and not characterized by clear directional dependence. The effects of spectral directionality were attributed to the characteristics of the subsoil only if they were consistently observed over the 46 min of recording. The frequency domain data were filtered with a triangular window to obtain a smoothing of 10%. The H/V spectral ratios of the microtremors were used to obtain stratified 1D models of the S velocity from each individual environmental vibration record. The inversion of the HVSR curves was performed using the relationship between the amplitude of the observed HV ratio and the ellipticity of the fundamental mode of the Rayleigh waves, assuming that the subsoil can be assimilated to a succession of homogeneous horizontal layers [3,16]. The uncertainty in the solution of the inverse problem was reduced by using the thickness and shear rate parameters of the inverse model related to the corresponding MASW as geometric and/or physical constraints [39,40]. The code used for the calculation of the synthetic H/V curves (Figure 4) is based on the simulation of the surface wave field (Rayleigh and Love waves) in a layered subsoil [23]. The inversion of the H/V curves was performed by choosing the closest MASW as the initial seismic velocity model.



**Figure 4.** H/V curves of the 80 HVSR measurements.

### 3.4.2. Cluster Analysis of the HVSR Peaks

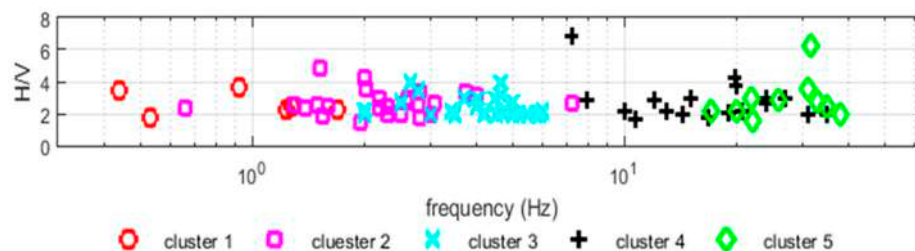
A modified non-hierarchical algorithm with K-means method [6] was applied to 80 HVSR measurements, acquired in the municipality of Enna. The proposed algorithm has the aim of reducing the units in  $k$  classes according to optimization criteria, minimizing the variance within the clusters and maximizing that between the clusters. This modified algorithm avoids any a priori choice of the operator which could influence the partition.

In order to verify the reliability and validity of the different partitions, the ratio between the variance between the clusters and the total variance of the data [6] was used:

$$R^2 = \frac{DEV_{OUT}}{DEV_T} \quad (1)$$



which made it possible to statistically choose a grouping of 5 clusters (Figure 5 and Table 1). This partitioning of the data was found to be compatible with the number of seismic discontinuities expected in the analyzed area.



**Figure 5.** Distribution in the frequency domain of the points of the 5 clusters.

**Table 1.** Minimum, average and maximum frequencies of the 5 clusters.

	Freq Min (Hz)	Freq Mean (Hz)	Freq Max (Hz)
cluster 1	0.44	1.02	1.70
cluster 2	0.66	2.46	7.20
cluster 3	2.00	4.14	6.00
cluster 4	7.20	18.19	35.00
cluster 5	17.00	27.58	38.00

#### 4. Results and Discussion

The analysis of the HVSR measurements, linked to the surface seismic velocity data obtained from the MASW measurements, made it possible to obtain the subsoil velocity models at greater depths. Figure 6 shows, for each of the clusters identified, a characteristic example of HVSR inversion, obtained by linking the velocity values of the more superficial layers with those obtained from the corresponding MASW model, together with a case without peaks, in which the corresponding MASW showed values of  $V_s$  characteristic of a sub-outcropping seismic bedrock.

By analyzing the inverse models obtained from each single measurement point, the values of the depth of the seismic bedrock were obtained, allowing the pattern of the surface of this subsoil layer to be reconstructed (Figure 7).

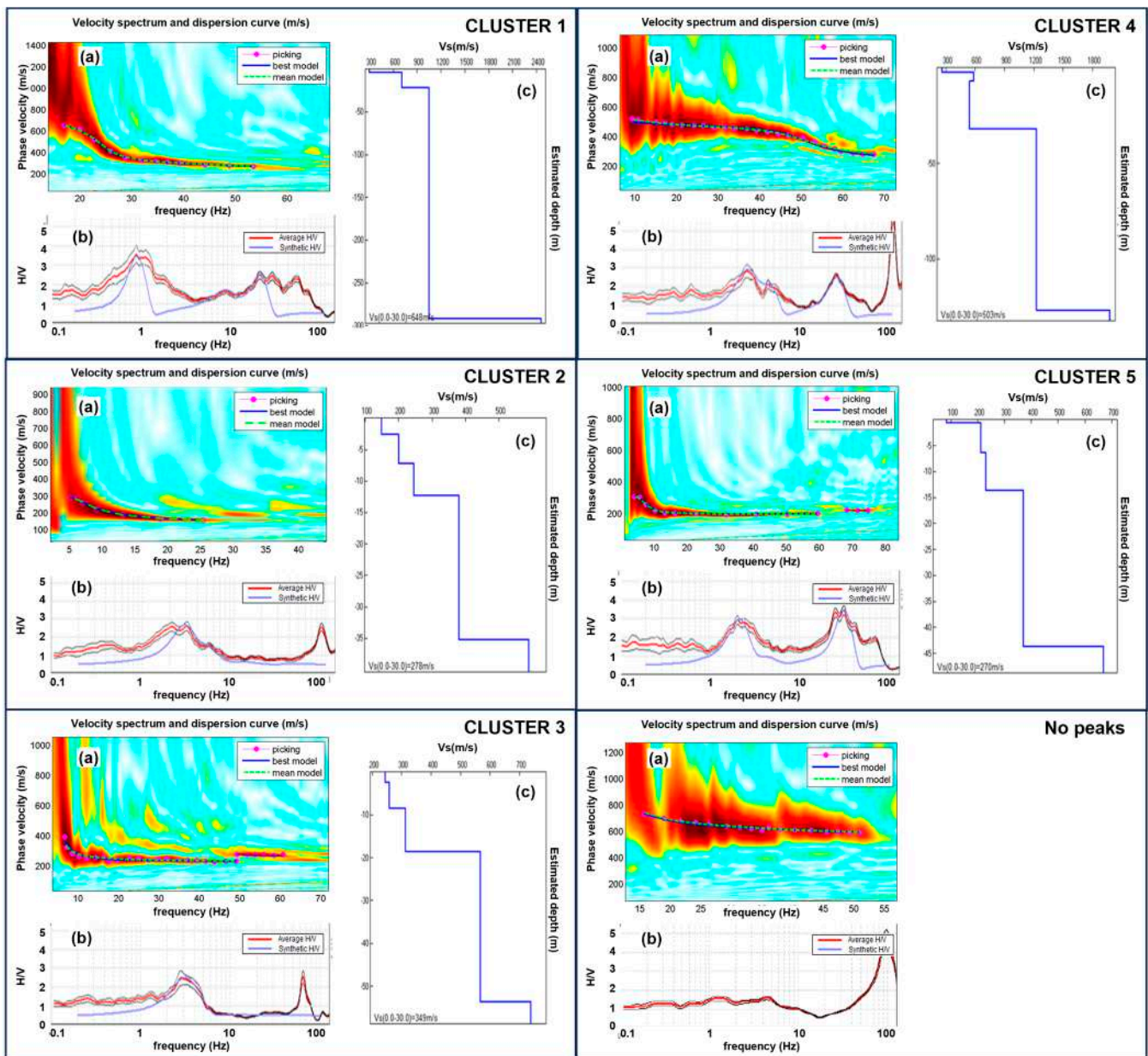
Furthermore, for each measurement point, we calculated the equivalent velocity of shear wave propagation  $V_{S,eq}$  defined as:

$$V_{S,eq} = H / \sum_{i=1}^N (h_i / V_{S,i}) \quad (2)$$

where  $h_i$  is the thickness of the  $i$ -th layer,  $V_{S,i}$  is the shear wave velocity in the  $i$ -th layer,  $N$  is the number of layers and  $H$  is the depth of the seismic bedrock, defined as that formation, made up of very rigid rock or soil, characterized by a  $V_s$  of not less than 800 m/s.

A map of the  $V_{S,eq}$  parameter was reconstructed for the territory (Figure 8) showing higher values of the parameter in the northern part of the territory and low values corresponding to the lower part, which corresponds to the central part of the territory.

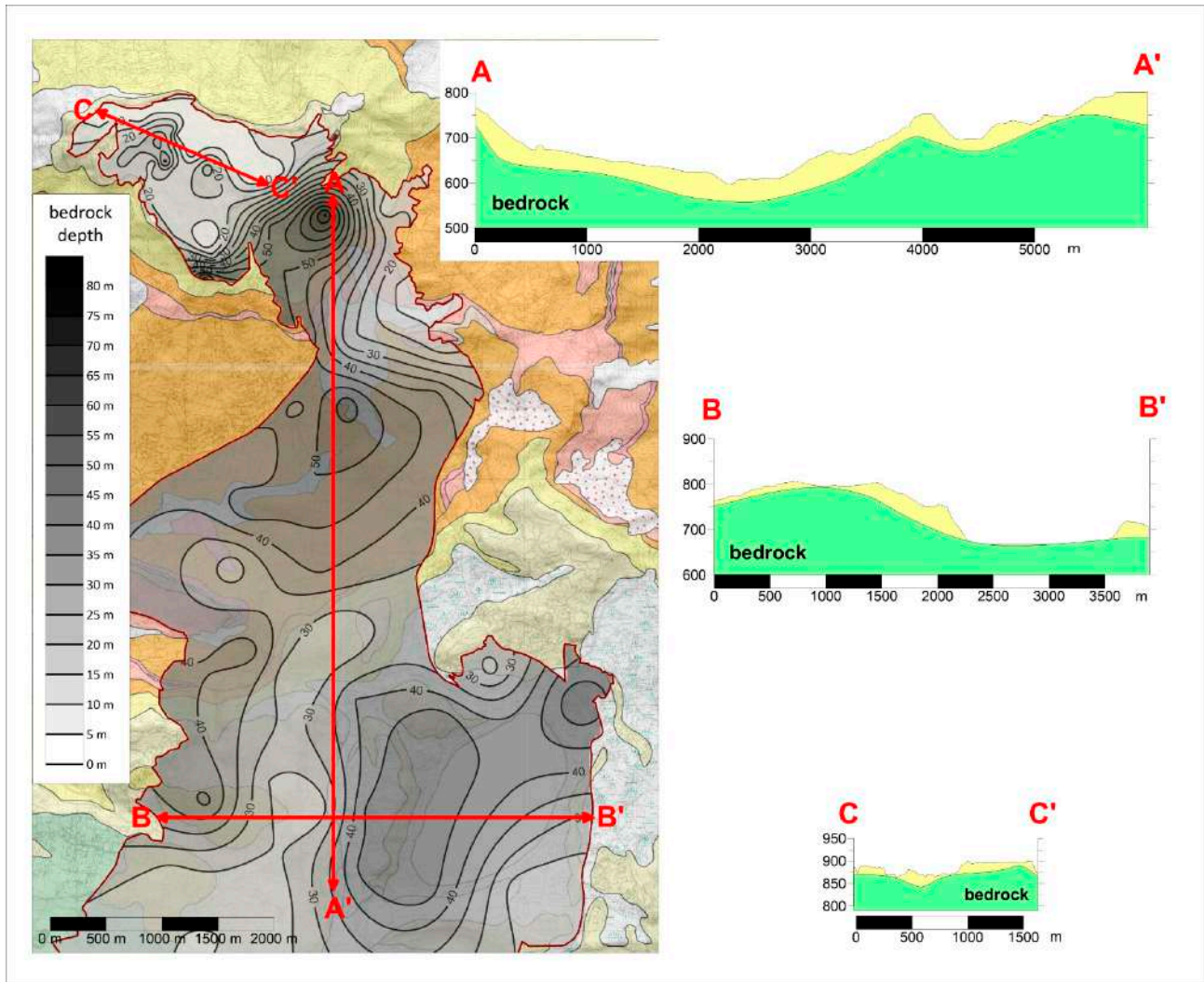
The site resonance frequencies grouped into five different clusters were attributed to different geological source. The stratigraphic interpretations were spatially correlated in order to reconstruct a three-dimensional geophysical model of the seismic discontinuities of the subsurface.



**Figure 6.** Examples of (a) MASW velocity spectrum and dispersion curve; (b) average (red) and synthetical (blue) HVSR curve; the gray lines indicate the standard deviation; (c) corresponding Vs model for each of the five clusters stated and in a case of no H/V peaks.

The results obtained were compared with the basic geology of the municipal area of Enna. The superimposition between the geological map and the five clusters (Figure 9) shows that, although the spatial limits of the clusters have not been constrained by the known stratigraphic contacts, their distribution seems to agree well, and the reconstructed seismic discontinuities show sub-horizontal geometries and compatible depths with those of the geological contacts present in the outcrop. In particular, cluster 1, characterized by mean frequencies of the centroid of 1.02 Hz and therefore related to a deep interface (greater than 50 m deep), seems to be related to the separation interface between the Enna Formation and the Trubi. Cluster 2, characterized by mean centroid frequencies of 4.14 Hz (less than 50 m deep), seems to be related to a discontinuity below or within the Terravecchia formation. Cluster 3, characterized by mean centroid frequencies of 18.19 Hz (less than 50 m deep), seems to be related to the separation interface between the two members of the Enna Formation, even if the cluster also includes some measurement points

external to this formation, which were excluded for the reconstruction of the separation surface, dividing the cluster into cluster 3a and cluster 3b. Cluster 4, characterized by mean centroid frequencies of 18.19 Hz (less than 18 m of depth), and cluster 5, characterized by mean centroid frequencies of 27.58 Hz (less than 13 m of depth), are probably due to seismic interfaces, which delimit the terrains of alteration.

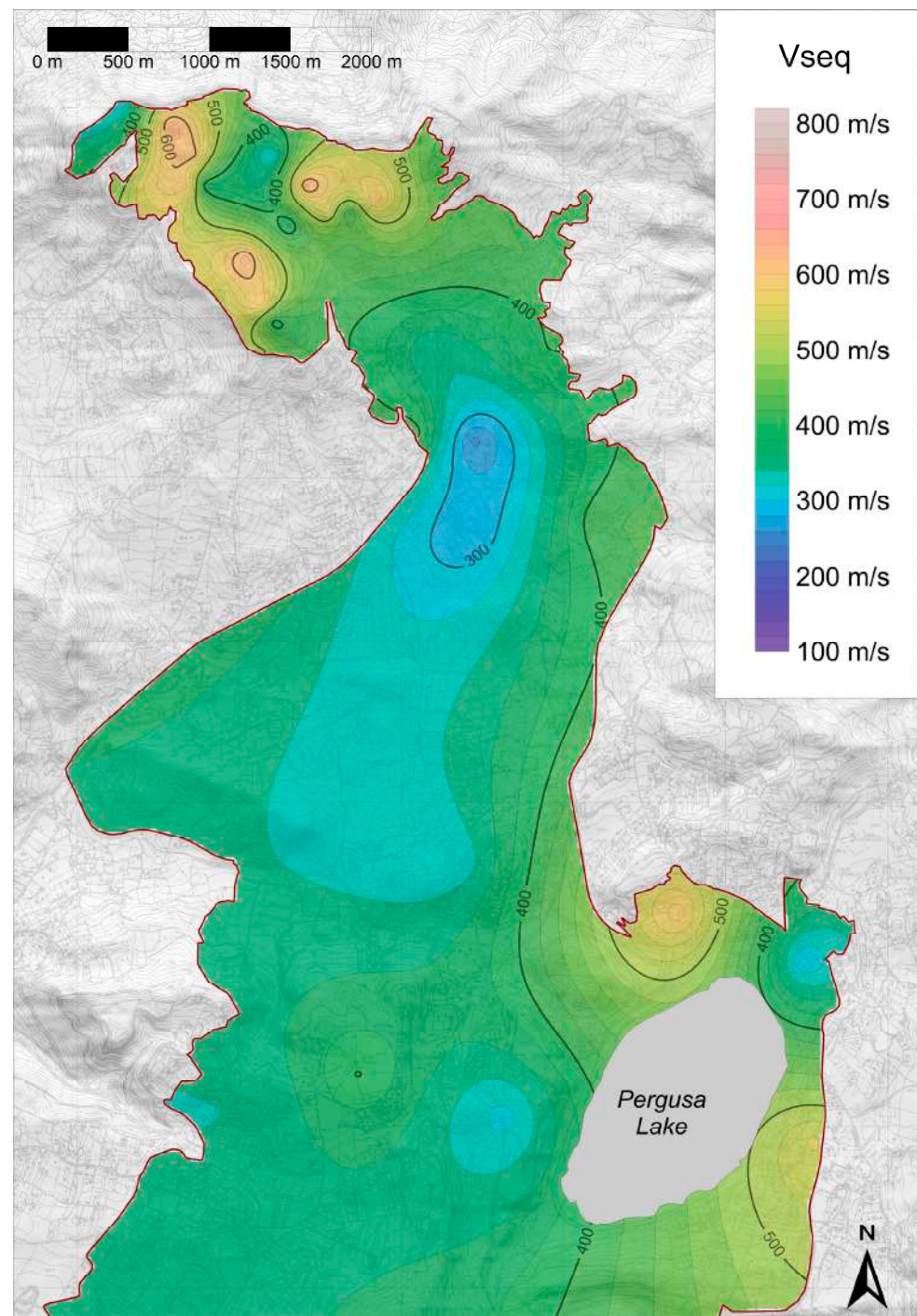


**Figure 7.** Seismic bedrock map. The brown line delimits the urbanized area, subject to seismic microzonation. Three sections show the behavior of the seismic bedrock in the subsurface.

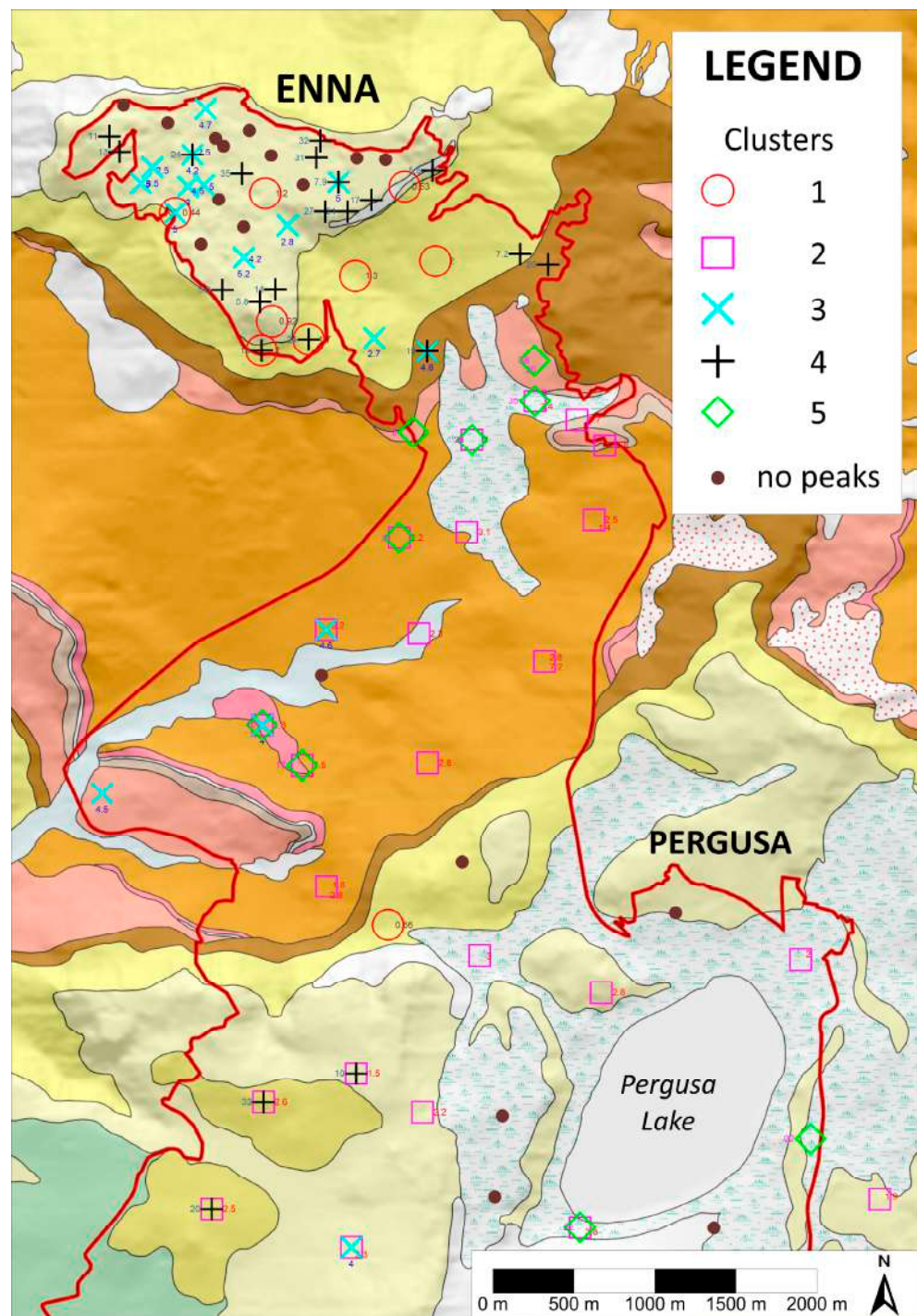
### 3D Modeling

Except for some measurements that did not show significant resonance peaks, the HVSR data partitioned into five clusters were inverted, assuming that the peaks relating to the same cluster are due to the same seismic generating surface. Therefore, five seismic surfaces were reconstructed, characterized by acoustic impedance contrasts such as being able to generate ground resonance frequency peaks. The surfaces obtained associated with each cluster are shown in Figure 10 together with the topography of the area.

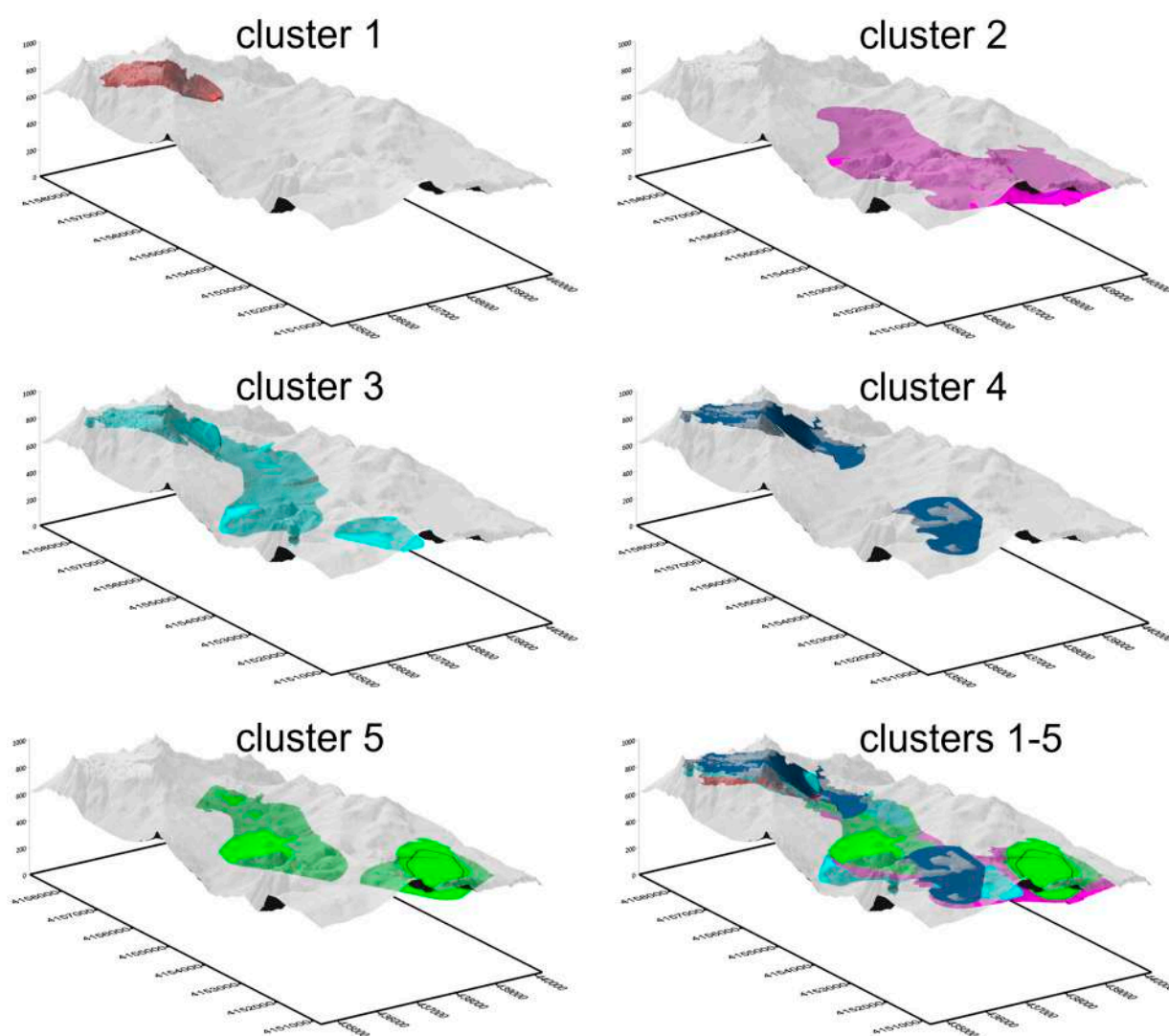
This type of evaluation can be decisive in the study of the seismic microzonation of an area to recognize the local geological and geomorphological conditions that could significantly alter the characteristics of the seismic shaking, generating stresses such as those producing permanent and critical effects.



**Figure 8.** Map of the value of  $V_{S,eq}$  of the sedimentary cover. The brown line delimits the urbanized area, subject to seismic microzonation.



**Figure 9.** Superimposition between the geological map and the five clusters. The red line delimits the urbanized area, subject to seismic microzonation.



**Figure 10.** 3D model of the reconstructed surfaces in relation to the topographic surface (in gray). The different color of each surface is that relating to the corresponding cluster (see legend of Figure 9).

## 5. Conclusions

The cluster analysis made it possible to divide the frequency peaks into five groupings, which were correlated to as many seismostratigraphic discontinuities of the subsoil. The inversion of the curves was then performed by extrapolating the depths of these surfaces at the measurement points relating to the cluster of that surface. This procedure made it possible to reconstruct the geometry and shape of these surfaces in the subsoil. However, it should be noted that some of these surfaces have been reconstructed with the few points falling within the relative cluster and that the resolution can therefore only be approximate and simplified. In fact, a reconstruction that is more consistent with reality requires numerous measurement points, especially for surfaces that show irregular trends. In the case of the municipal area of Enna, the trend of the surfaces appears to be sub-horizontal in many cases, as confirmed by the surface geology, so the reconstruction is realistic.

The application of cluster analysis to the analysis of the frequency peaks of the HVSR measurements has therefore shown good potential in the reconstruction of seismic and geological models of the subsoil, especially in the case of geometries that are simple to reconstruct.

**Author Contributions:** Conceptualization, P.C. and R.M.; methodology, P.C. and R.M.; validation, P.C. and R.M.; formal analysis, P.C.; investigation P.C. and R.M.; data curation, P.C. and R.M.; writing—original draft preparation, P.C. and R.M.; writing—review and editing, P.C. and R.M.; visualization, R.M.; supervision, P.C. and R.M.; project administration, R.M.; funding acquisition, R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by PRISMA Project (Piattaforme cloud interoperabili per smart government): PON04a2\_A.

**Data Availability Statement:** Data can be obtained upon request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Nakamura, Y. Method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. *Q. Rep. RTRI (Railw. Tech. Res. Inst.)* **1989**, *30*, 25–33.
2. Nakamura, Y. Clear Identification of Fundamental Idea of Nakamura’s Technique and its Applications. In Proceedings of the 12th World Conference on Earthquake Engineering, Auckland, New Zealand, 30 January–4 February 2000; Volume 2656.
3. Fäh, D.; Kind, F.; Giardini, D. Inversion of local S wave velocity structures from average H/V ratios, and their use for the estimation of site-effects. *J. Seismol.* **2003**, *7*, 449–467. [[CrossRef](#)]
4. Parolai, S.; Bindi, D.; Augliera, P. Application of the Generalized Inversion Technique (GIT) to a microzonation study: Numerical simulations and comparison with different site-estimation techniques. *Bull. Seismol. Soc. Am.* **2000**, *90*, 286–297. [[CrossRef](#)]
5. Martorana, R.; Agate, M.; Capizzi, P.; Cavera, F.; D’Alessandro, A. Seismo-stratigraphic model of “La Bandita” area (Palermo Plain, Sicily) through HVSR inversion constrained by stratigraphic data. *Ital. J. Geosci.* **2018**, *137*, 73–86. [[CrossRef](#)]
6. Capizzi, P.; Martorana, R. Analysis of HVSR Data Using a Modified Centroid-Based Algorithm for Near-Surface Geological Reconstruction. *Geosciences* **2022**, *12*, 147. [[CrossRef](#)]
7. Hartigan, J.A. *Clustering Algorithms*; Wiley: New York, NY, USA, 1975.
8. Adelfio, G.; Chiodi, M.; D’Alessandro, A.; Luzio, D.; D’Anna, G.; Mangano, G. Simultaneous seismic wave clustering and registration. *Comput. Geosci.* **2012**, *44*, 60–69. [[CrossRef](#)]
9. D’Alessandro, A.; Mangano, G.; D’Anna, G.; Luzio, D. Waveforms clustering and single-station location of microearthquake multiplets recorded in the northern Sicilian offshore region. *Geophys. J. Int.* **2013**, *194*, 1789–1809. [[CrossRef](#)]
10. D’Alessandro, A.; Luzio, D.; Martorana, R.; Capizzi, P. Selection of time windows in the Horizontal to Vertical Noise Spectral Ratio by means of cluster analysis. *Bull. Seismol. Soc. Am.* **2016**, *106*, 560–574. [[CrossRef](#)]
11. Martorana, R.; Capizzi, P.; Avellone, G.; Siragusa, R.; D’Alessandro, A.; Luzio, D. Assessment of a geological model by surface wave analyses. *J. Geophys. Eng.* **2017**, *14*, 159–172. [[CrossRef](#)]
12. Martorana, R.; Capizzi, P.; D’Alessandro, A.; Luzio, D.; Di Stefano, P.; Renda, P.; Zarcone, G. Contribution of HVSR measures for seismic microzonation studies. *Ann. Geophys.* **2018**, *61*, SE225. [[CrossRef](#)]
13. Capizzi, P.; Martorana, R.; Stassi, G.; D’Alessandro, A.; Luzio, D. Centroid-based cluster analysis of HVSR data for seismic microzonation. In Proceedings of Near Surface Geoscience 2014—20th European Meeting of Environmental and Engineering Geophysics, Athens, Greece, 14–18 September 2014. [[CrossRef](#)]
14. Picotti, S.; Francese, R.; Giorgi, M.; Pettenati, F.; Carcione, J.M. Estimation of glaciers thicknesses and basal properties using the horizontal-to-vertical component spectral ratio (HVSR) technique from passive seismic data. *J. Glaciol.* **2017**, *63*, 229–248. [[CrossRef](#)]
15. Scherbaum, F.; Hinzen, K.-G.; Ohrnberger, M. Determination of shallow shear-wave velocity profiles in Cologne, Germany area using ambient vibrations. *Geophys. J. Int.* **2003**, *152*, 597–612. [[CrossRef](#)]
16. Arai, H.; Tokimatsu, K. S-wave velocity profiling by inversion of microtremor H/V spectrum. *Bull. Seismol. Soc. Am.* **2004**, *94*, 53–63. [[CrossRef](#)]
17. Parolai, S.; Picozzi, M.; Richwalski, S.M.; Milkereit, C. Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes. *Geophys. Res. Lett.* **2005**, *32*, L01303. [[CrossRef](#)]
18. Picozzi, M.; Parolai, S.; Richwalski, S.M. Joint inversion of H/V ratios and dispersion curves from seismic noise: Estimating the S-wave velocity of bedrock. *Geophys. Res. Lett.* **2005**, *32*, L11308. [[CrossRef](#)]
19. Imposa, S.; Grassi, S.; De Guidi, G.; Battaglia, F.; Lanaia, G.; Scudero, S. 3D subsoil model of the San Biagio ‘Salinelle’ mud volcanoes (Belpasso, SICILY) derived from geophysical surveys. *Surv. Geophys.* **2016**, *37*, 1117–1138. [[CrossRef](#)]
20. Imposa, S.; Grassi, S.; Fazio, F.; Rannisi, G.; Cino, P. Geophysical surveys to study a landslide body (north-eastern Sicily). *Nat. Hazards* **2017**, *86*, 327–343. [[CrossRef](#)]
21. Zor, E.; Özalaybey, S.; Karaaslan, A.; Tapirdamaz, M.C.; Özalaybey, S.Ç.; Tarancioglu, A.; Erkan, B. Shear wave velocity structure of the Izmit Bay area (Turkey) estimated from active-passive array surface wave and single-station microtremor methods. *Geophys. J. Int.* **2010**, *182*, 1603–1618. [[CrossRef](#)]
22. Capizzi, P.; Martorana, R. Integration of constrained electrical and seismic tomographies to study the landslide affecting the Cathedral of Agrigento. *J. Geophys. Eng.* **2014**, *11*, 045009. [[CrossRef](#)]

23. Castellaro, S. The complementarity of H/V and dispersion curves. *Geophysics* **2016**, *81*, T323–T338. [[CrossRef](#)]
24. Panzera, F.; Sicali, S.; Lombardo, G.; Imposa, S.; Gresta, S.; D’Amico, S. A microtremor survey to define the subsoil structure in a mud volcanoes area: The case study of Salinelle (Mt. Etna, Italy). *Environ. Earth Sci.* **2016**, *75*, 1140. [[CrossRef](#)]
25. Tortorici, L.; Grasso, M.; Catalano, S.; Monaco, C. Note Illustrative della Carta Geologica d’Italia Alla Scala 1:50.000, Foglio n. 631 “Caltanissetta”. Regione Siciliana-Ispira. 2014. Available online: [https://www.isprambiente.gov.it/Media/carg/note\\_illustrative/631\\_Caltanissetta.pdf](https://www.isprambiente.gov.it/Media/carg/note_illustrative/631_Caltanissetta.pdf) (accessed on 1 June 2023).
26. Molnar, S.; Cassidy, J.F.; Castellaro, S.; Cornou, C.; Crow, H.; Hunter, J.A.; Matsushima, S.; Sánchez-Sesma, F.J.; Yong, A. Application of Microtremor Horizontal-to-Vertical Spectral Ratio (MHVSR) Analysis for Site Characterization: State of the Art. *Surv. Geophys.* **2018**, *39*, 613–631. [[CrossRef](#)]
27. Molnar, S.; Sirohey, A.; Assaf, J.; Bard, P.-Y.; Castellaro, S.; Cornou, C.; Cox, B.; Guillier, B.; Hassani, B.; Kawase, H.; et al. A review of the microtremor horizontal-to-vertical spectral ratio (MHVSR) method. *J. Seismol.* **2022**, *26*, 653–685. [[CrossRef](#)]
28. Field, E.H.; Jacob, K. The theoretical response of sedimentary layers to ambient seismic noise. *Geophys. Res. Lett.* **1993**, *20*, 2925–2928. [[CrossRef](#)]
29. Lermo, J.; Chávez-García, F.J. Site effect evaluation using spectral ratios with only one station. *Bull. Seismol. Soc. Am.* **1993**, *83*, 1574–1594. [[CrossRef](#)]
30. Mucciarelli, M. Reliability and applicability of Nakamura’s technique using microtremors: An experimental approach. *J. Earthq. Eng.* **1998**, *2*, 625–638. [[CrossRef](#)]
31. Mucciarelli, M.; Gallipoli, M.R.; Arcieri, M. The stability of the horizontal-to-vertical spectral ratio of triggered noise and earthquake recordings. *Bull. Seismol. Soc. Am.* **2003**, *93*, 1407–1412. [[CrossRef](#)]
32. Lachet, C.; Bard, P.Y. Numerical and theoretical investigations on the possibilities and limitations of Nakamura’s technique. *J. Phys. Earth* **1994**, *42*, 377–397. [[CrossRef](#)]
33. Konno, K.; Ohmachi, T. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. *Bull. Seismol. Soc. Am.* **1998**, *88*, 228–241. [[CrossRef](#)]
34. Fäh, D.; Kind, F.; Giardini, D. A theoretical investigation of average H/V ratios. *Geophys. J. Int.* **2001**, *145*, 535–549. [[CrossRef](#)]
35. SESAME Project. Guidelines for the Implementation of the H/V Spectral Ratio Technique on Ambient Vibrations. Measurements, Processing and Interpretation, SESAME European Research Project WP12–Deliverable D23.12, December 2004. Available online: [http://sesame.geopsy.org/Papers/HV\\_User\\_Guidelines.pdf](http://sesame.geopsy.org/Papers/HV_User_Guidelines.pdf) (accessed on 1 June 2023).
36. Horike, M.; Zhao, B.; Kawase, H. Comparison of site response characteristics inferred from microtremors and earthquake shear waves. *Bull. Seismol. Soc. Am.* **2001**, *81*, 1526–1536. [[CrossRef](#)]
37. Lunedei, E.; Albarello, D. On the seismic noise wavefield in a weakly dissipative layered Earth. *Geophys. J. Int.* **2009**, *177*, 1001–1014. [[CrossRef](#)]
38. Bonnefoy-Claudet, S.; Cotton, F.; Bard, P.-Y. The nature of noise wavefield and its applications for site effects studies. A literature review. *Earth Sci. Rev.* **2006**, *79*, 205–227. [[CrossRef](#)]
39. Parolai, S.; Richwalski, S.M.; Milkereit, C.; Fäh, D. S-wave velocity profiles for earthquake engineering purposes for the Cologne Area (Germany). *Bull. Earthq. Eng.* **2006**, *4*, 65–94. [[CrossRef](#)]
40. Castellaro, S.; Mulargia, F. VS30 estimates using constrained H/V measurements. *Bull. Seismol. Soc. Am.* **2009**, *99*, 761–773. [[CrossRef](#)]
41. Mi, B.; Hu, Y.; Xia, J.; Socco, L.V. Estimation of horizontal-to-vertical spectral ratios (ellipticity) of Rayleigh waves from multistation active-seismic records. *Geophysics* **2019**, *84*, EN81–EN92. [[CrossRef](#)]
42. Mi, B.; Xia, J. Inversion of Rayleigh-wave ellipticity for shallow strong-velocity-contrast structures. *Geophysics* **2023**, *88*, KS89–KS100. [[CrossRef](#)]
43. Park, C.B.; Miller, R.D.; Xia, J. Multichannel analysis of surface waves. *Geophysics* **1999**, *64*, 800–808. [[CrossRef](#)]
44. Fabbris, L. *Analisi Esplorativa di Dati Multidimensionali*; Cleup Editore: Padova, Italy, 1983.
45. Barbarito, L. *L’analisi di Settore: Metodologia e Applicazioni*; Franco Angeli: Milano, Italy, 1999.
46. Gan, G.; Ma, C.; Wu, J. *Data Clustering: Theory, Algorithms, and Applications*; Cambridge University Press: Cambridge, UK, 2007; p. 184, ISBN 9780898716238.
47. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis, Wiley Series in Probability and Statistics*, 5th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2011; p. 332.
48. Rand, W.M. Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **1971**, *66*, 846–850. [[CrossRef](#)]
49. Ohsumi, N. Evaluation procedure of agglomerative hierarchical clustering methods by fuzzy relations. In *Data Analysis and Informatics: Proceedings of the II international Symposium on Data Analysis and Informatics*; Diday, E., Ed.; North Holland, Edgell and Geisler: Amsterdam, The Netherlands, 1980.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.