Application of a method to determine S and P wave velocities from surface waves data analysis in presence of sharp lateral variations

APPLICATION OF A METHOD TO DETERMINE S AND P WAVE VELOCITIES FROM SURFACE WAVES DATA ANALYSIS IN PRESENCE OF SHARP LATERAL VARIATIONS

F. Khosro Anjom[1], A. Arabi[1], L.V. Socco[1], C. Comina[2]
[1] *DIATI (Dipartimentoo di Ingegneria dell'Ambiente, del territorio e delle infrastrutture) Politecnico di Torino, Italy*
[2] *Dipartimento di Scienze della Terra, Università di Torino, Italy*

Introduction. Seismic surface waves are becoming a leading method to estimate S-wave velocity (VS) through dispersion curve (DC) extraction and inversion (Socco *et al*., 2010). Socco *et al*. (2017) showed that a relationship between wavelength and depth (W/D) of surface waves can be determined and used to directly transform the DC into the time-average VS model, avoiding the inversion process. Surface waves are usually considered unsuitable for estimating P-wave velocity (VP), or Poisson's ratio, due to the low sensitivity of dispersion curve to these parameters (Nazarian,1984). Socco and Comina (2017) found that this W/D relationship is sensitive to Poisson's ratio. A method was therefore developed to determine also the time-average VP by using the W/D relationship. These methods are robust in terms of accuracy and efficient in terms of computational cost. They can be applied to a set of dispersion curves over a seismic line to provide a pseudo 2D time-average VS and VP model along the line. The methods require a reference DC and its associated VS profile to get the W/D relationship. However, using a unique reference curve for the whole line becomes problematic when sharp lateral variations or zones with different velocities are present.

In this work, a hierarchical clustering algorithm is developed to select ensembles of DCs within a uniform zone that can be interpreted using the same reference DC. Then, reference W/D relationships are estimated in each uniform zone and combined with DCs to determine time-average VS and VP models over the whole area. In the following, the clustering method is first briefly explained, then, the application of the method to field data from a test site (CNR in Turin) is shown and discussed.

Methodology. A hierarchical clustering method has been used to group the DCs. In hierarchical clustering, there is no need to set a predefined number of clusters, this is a great advantage since we want a method able to identify uniform zones without knowing a priori the expected lateral variability of the area. The result of hierarchical clustering is a dendrogram which shows the nested grouping of observations and similarity levels at which grouping change (Maimon and Rokach, 2000). Hierarchical methods can be divided into two subgroups (Maimon and Rokach, 2000): i) agglomerative - a bottom up approach in which every observation is considered a cluster and similar clusters are identified and merged together to make bigger clusters; ii) divisive - a top down approach in which the observations all start in a single cluster and at each step they are removed from the cluster to create a new cluster or join other clusters.

Tests on synthetic and real data showed that agglomerative and divisive methods have similar performances for our purpose. We selected agglomerative clustering because it makes the process tracking easier. A measure of dissimilarity is needed to merge the clusters. This is usually achieved by appropriate metric and linkage criterion. Euclidean distance is used here to determine the metric between two dispersion curves:

$$D(v_i, v_j) = \sqrt{(v_{i_1} - v_{j_1})^2 + (v_{i_2} - v_{j_2})^2 + \cdots + (v_{i_p} - v_{j_p})^2} \quad (1)$$

where D is the distance, v_i and v_j are phase velocity of the two dispersion curves at different frequencies. Linkage criterion determines the distance between sets of observations or clusters. Average linkage, in which the distance between clusters is calculated based on the average distance between each component of one cluster to each component of the other cluster, is used here to maximize the contribution of all DCs in the clusters.

After the DCs are grouped, in each group (cluster) the DC with the highest quality and broad

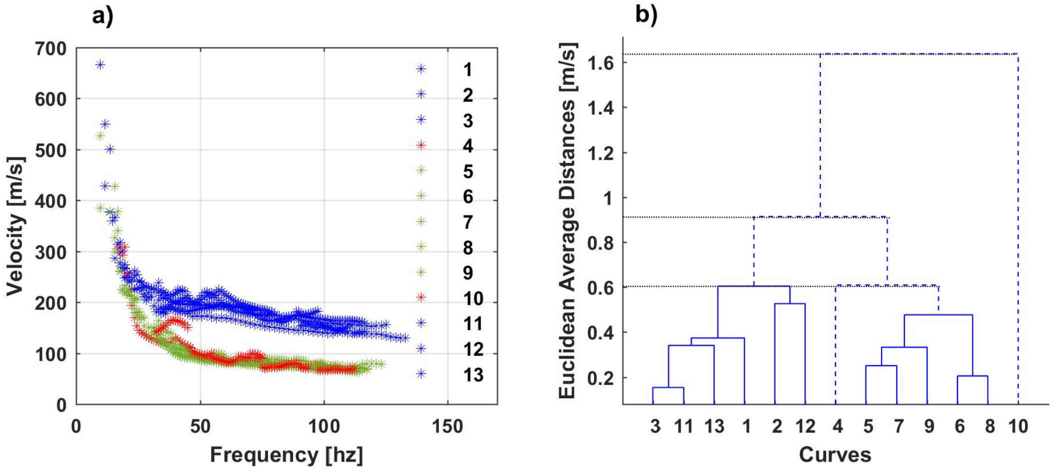

Fig. 1 - a) the dispersion curves along the seismic line from CNR dataset b) the dendrogram showing the cluster system of the CNR's dispersion curves.

frequency band is selected as reference and inverted using a Monte Carlo Inversion (MCI) (Socco and Boiero, 2008). In the inversion, layer thickness, S-wave velocity and Poisson's ratio are randomly sampled from wide uniform distributions. This allows a VS model to be estimated without bias on Poisson's ratio values which is a prerequisite for the adopted methodology. The reference W/D relationship is then computed and used to determine the time-average VS of all profiles in the cluster. By exploiting the W/D relationship sensitivity to Poisson's ratio the time-average VP of the whole cluster is then determined.

Results. The data have been acquired during a student-team work at a test site located at CNR headquarter in Torino. A squared area (5x5m) was excavated and filled with uniform non-compacted sand at the site. The bottom of the sand volume goes down to 2.5 m depth. A seismic line was acquired with 72, 4.5 Hz vertical geophones with 0.3 m spacing and with multiple shots by a hammer source. The data were recorded using three Geode modules with 512 ms record duration and 0.125 ms sampling interval. The data were processed in f-*k* domain using a moving window to estimate 13 DCs along the seismic line (Teodor *et al.*, 2017). Fig. 1a shows the DCs numbered according to their position along the line and with different colors to highlight curves located inside the sand (green), outside the sand (blue) or close to the sand boundaries (red). The clustering method produced the dendrogram reported in Fig. 1b. The horizontal axis of the dendrogram reports the DC numbers and the vertical axis reports the distance between linked clusters. The higher the vertical nodes the further the clusters are from each other.

The dendrogram shows two major clusters corresponding to the curves that are inside and outside the sand body. For the two curves close to the boundaries it shows that curve 10 could be considered as a single cluster (outlier). Curve 4 is connected to its cluster with a very high distance. Curves 4 and 10 are then considered outliers that should be processed individually and they were not included in the velocity model estimation. A DC from each cluster was used to estimate the VS model through MCI and to obtain the reference W/D. The application of the W/D to the other DCs of the cluster provided the time-average VS models reported in Fig. 2a. The Poisson's ratio estimate from the W/D relationship of each cluster was then applied to the VS model to transform it into a time-average VP model shown in Fig. 2b.

Estimated time-average VS and VP clearly show a low velocity zone in the middle of the profile related to the presence of the sand body. This low velocity zone is partially laterally continued outside of the sand body in the left part of the sand body. This is due to the limited

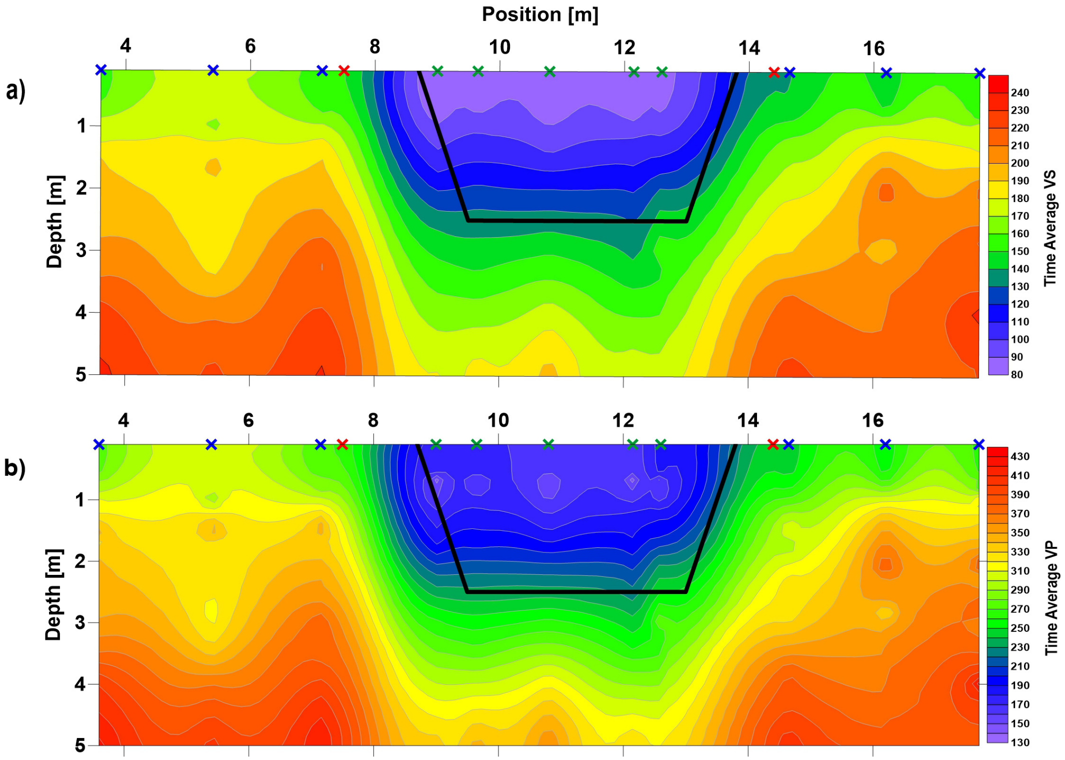

Fig. 2 - a) The estimated time-average VS model from the CNR data. b) The estimated time-average VP model from the CNR data; cross section of the sand body is displayed by the black line and locations of the different DCs are shown by crosses whose color indicates the cluster.

number of DCs and large distance between DCs near the sand body border that create low lateral resolution and consequent artifact due to interpolation. This could be resolved by retrieving more DCs during processing stage. The comparison of the velocities with those obtained by traditional DC inversion and P-wave tomography show variations mostly under 10%.

Conclusion. We have shown that time-average VS and VP can be estimated, even in presence of sharp lateral variations, by directly transforming DCs using Socco *et al.* (2017) and Socco and Comina (2017). The results of the time-average VS and VP in Fig. 2 shows that the methods can outline, with great level of confidence, the near surface complexity and they could be a powerful alternative in near surface characterizations.

The developed hierarchical clustering method grouped the DCs into two main groups outside and inside the sand body. Even though the DCs from CNR are not particularly challenging in terms of clustering, since two clear groups emerges by eye, still the developed clustering algorithm resulted to give very good information with respect to outliers. This could be potentially more useful over wider datasets and more complicated subsurface areas.

Acknowledgments The authors would like to thank TOTAL E&P-RECHERCHE DEVELOPMENT for supporting the “Novel geophysical methods for retrieving near surface P-and S-wave velocity models” research project.

References

Nazarian, S., 1984, In situ determination of elastic moduli of soil deposits and pavement systems by spectral-analysis-of-surface waves method: Ph.D. dissertation, The University of Texas at Austin.

Oded Maimon and Lior Rokach. 2010. *Data Mining and Knowledge Discovery Handbook* (2nd ed.). Springer Publishing Company, Incorporated. doi: 10.1007/978-0-387-09823-4

Socco, L. V., C. Comina, and F. Khosro Anjom, 2017, Time-average velocity estimation through surface-wave analysis: Part 1 — S-wave velocity: Geophysics, 82(3), U49-U59.

Socco, L. V. and C. Comina, 2017, Time-average velocity estimation through surface-wave analysis: Part 2 — P-wave velocity: Geophysics, 82(3), U61-U73.
Socco, L. V., S. Foti, and D. Boiero, 2010, Surface wave analysis for building near surface velocity models: established approaches and new perspectives: Geophysics, 75(5), 75A83–75A102, doi: 10.1190/1.3479491.
Socco, L. V., Boiero, Daniele, 2008, Improved Monte Carlo inversion of surface wave data: Geophysical Prospecting. 56(3), doi: 10.1111/j.1365-2478.2007.00678.x
Teodor, D., Comina, C., Socco, L.V. , Brossier, R., Trinh P. T. & Virieux, J. (2017) - Initial Model Design for Full-Waveform Inversion - Preliminary Elastic Modelling from Surface Waves Data Analysis. In. 36th GNGTS Nat. Conv.