# Automatic mapping of the velocity field in seismic reflection data using unsupervised learning K-means clustering

Marcos Augusto Lima da Luz*, CPGG/UFBA, Michelângelo Gomes da Silva, CPGG/UFBA, Milton José Porsani, CPGG/UFBA.



## Abstract

Reflection seismic is one of the most important geophysical methods, being largely responsible for the numerous discoveries of current oil reserves. Its principle is based on the emission and reception of mechanical waves from artificial sources. Despite the method presenting a good resolution of the geology, the seismic data always present noise and distortions that hinder the interpretation of the studied geological features, a reason that requires attention during the processing of these data. The noises can be attenuated with adequate processing sequences, already the distortions, such as those caused by low-velocity zones (LVZ), are more difficult to remove from seismic sections. In the ZBV regions, occurs a reduction in the velocity of the seismic wave, which can lead to the appearance of false geological structures (artifacts) or even hide true structures. Thus, a process that does not consider this phenomenon could lead to a false interpretation of the data. To eliminate the change in time of the reflections, the static correction technique is used, where the correct determination of the velocity field is a prime factor for its success. Thus, this work has as its objective to use the unsupervised learning K-means clustering technique for more accurate and automatic identification of the velocity field of primary reflections in synthetic and real seismic data. With the advancement of artificial intelligence, many techniques could be adapted for different applications, we use the K-means clustering that makes it possible to partition a dataset into K distinct and non-overlapping clusters, that is, each observation belongs to at least one of the K clusters and no observation belongs to more than one cluster. These characteristics are important when estimating the real allocation of a variable. In ZBV regions, the signal suffers great attenuation, demonstrating atypical behavior in relation to adjacent signals. The application of K-means clustering on these data, classifies the variables within different groups, so that the intra-cluster variation is minimized by the sum of squares of the Euclidean distance between the variables and their centroids, and performs updates by calculating the new mean of the values within your defined cluster. This process manages to coherently reallocate the values of the variables in the groups with the greatest similarity, such as, for example, the real values of the speed of the signals arising from these regions. Several tests were performed on synthetic seismic data with different levels of complexity (including the Marmousi seismic data) and real data such as the Gulf of Mexico (this one for having a large number of studies). The results were satisfactory, returning levels of 98% of similarity with the already known values of the velocity field in the synthetic data and 95% in the real Gulf of Mexico data when compared with the values obtained by the traditional picking method. In this way, it is possible to perceive that the technique allows the automatic obtaining of the velocity field in seismic reflection data, with a high level of precision and without the need to carry out manual picking.

Keywords: Reflection Seismic, Artificial Intelligence, Clustering, K-means.