Geophysical Research Abstracts
Vol. 12, EGU2010-13066, 2010
EGU General Assembly 2010

Parallel Grid approach to solve Feature Selection problem in volcanic infrasound signals classification

Danilo Reitano (1), Fabrizio Pistagna (2), Gaetano Russo (3), and Vincenzo Valenti (4)

(1) (1) Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, P.zza Roma, 2 – 95125, Catania Italy (reitano@ct.ingv.it), (2) Istituto Nazionale di Fisica Nucleare, INFN, Via S. Sofia , 64 – 95123 Catania Italy, (3) Consorzio Cometa, Via S. Sofia , 64 – 95123 Catania Italy, now After Sales engineer at Elica SpA, (4) Università degli studi di Catania, Facoltà di Ingegneria, DIIT, Via S. Sofia , 64 – 95123 Catania Italy

The continuous monitoring of an active volcano, such as Mt. Etna (Sicily, Italy), represents one of the main tasks for the Italian National Institute of Geophysics and Volcanology (INGV), Catania Branch. Around the volcano summit area, four infrasound sensors have been installed during the last recent years, which allow to acquire, real time waveforms that are subsequently stored on a server, located inside the INGV Control Room. A new method here introduced, based on Genetic Algorithms (GA), has been used to analyze the data coming from the remote infrasound sensors stations. In particular, the acquired signals have been processed by a custom modular software: the first module allows the visual manipulation, filtering and, in order to optimize performances, resampling the data to better elaborate them. The second module, using an alghorithm (G. Russo, 2009) based on two different thresholds (upper and lower) and the standard deviation, is able to recognize and collect infrasound events (IE) from the stored data. In the third step, the Green & Nueberg algorithm (2006) is used to correlate different families of IE and define the clusters nodes. Once a minimum number of families are characterized, they define the main features inside each cluster. Feature extraction process is a very complex algorithm due to the large number of requested correlations. In order to speed up the time needed to carry out so many simulations, the code has been deployed and executed on the Sicilian Grid infrastructure owned and managed by the Consorzio Cometa, a not-for-profit organisation including INGV among its members. The infrastructure, distributed across the Sicilian territory, is composed of 7 sites for a total of about 2000 CPU cores and more than 250 TB of storage. All the sites of the infrastructure are equipped with low latency Infiniband networks and are installed with MPI libraries. A complete workflow has been created from scratch to execute the various phases of the modular method described above,. It includes a parallel Matlab™ portion of code to solve the first module of the method, while an Artificial Neural Network (ANN), written in ANSI-C, extending the PGAPack libraries and using OpenMPI libraries for the parallel computation of the Genetic Algorithms, has been used for the last phase of the method. Final results as well future directions will be presented and discussed.