A clustering analysis of three dimensional ground deformation map estimated by integrating DInSAR and GPS dataset

G. Nunnari, G. Puglisi, A. Spata, F. Guglielmino, A. Bonforte, P. Montalto

Abstract

In this paper a clustering analysis based on the combination of the Self-Organizing Map (SOM) and the K-means method is applied to three dimensional ground deformation map obtained by integrating sparse Global Positioning System (GPS) and Differential Interferometric Synthetic Aperture Radar (DInSAR) acquired at Mt Etna in the period 2003-2004. This analysis is aimed to partition the whole displacement field into subsets sharing some common displacements features in order to recognize and classify deformation patterns affecting different sectors of Etna volcano. Results have been also confirmed by a fuzzy c-mean analysis.

Introduction

The identification of ground deformation movements and the characterization of active faults are priority targets for the geophysical monitoring of Mt Etna, the most active volcano in Europe. In this framework both DInSAR and GPS techniques are successfully used to monitor ground deformation at Mt Etna [1-4]. In order to take advantage of the complementary nature of satellite and geodetic data, current efforts of the scientific community are devoted to develop suitable algorithms able to efficiently integrate these data. Indeed although satellite DInsar enables studying ground deformations with a spatial resolution unprecedented by any other geodetic techniques, it is characterized by a low temporal resolution and provides a mono-dimensional measurement of deformations. On the other hand although GPS is the most suitable technique for measuring ground deformation with sub-cm accuracy level, it provide a point wise 3D displacement vector referring to the specific geodetic benchmark where the antenna is set up; consequently, the spatial resolution of the measurement of the ground deformations is depending from the network geometry and thus is usually low.

Here we present an approach to identify deformation patterns based on the joint use of a recently proposed technique to combine DInSAR data and GPS measurements referred to as SISTEM (Guglielmino et al 2009?), and the Self-Organizing Map (SOM).

DInSAR and GPS integration (SISTEM method)

Let assume that a geodynamic process (e.g. intrusions of magma or earthquakes) deforms a portion of Earth's surface; under the hypothesis of infinitesimal and homogeneous strain we define an arbitrary point P, having position $x_0=(x_{10}, x_{20}, x_{30})$, and N surrounding experimental points (EPs) whose positions and displacements are respectively $\boldsymbol{x}_{(n)}=(x_{1(n)}, x_{2(n)}, x_{3(n)})$ and $\boldsymbol{u}_{(n)}=(u_{1(n)}, u_{2(n)}, u_{3(n)})$ where $n=1..N$. Is a such hypothesis, adopting a linear approach, the problem of estimating the displacement components U_i (i=1..3) of the point P, from the experimental data $\boldsymbol{u}_{(n)}=(u_{1(n)}, u_{2(n)}, u_{3(n)})$, can be modelled by the N equations [6]:

$$u_{i(n)}(x) = H_{ij}\Delta x_{j(n)} + U_i \quad (i, j = 1..3) \qquad (1)$$

where $\Delta x_{j(n)}=x_{j(n)}-x_{j0}$ are the relative positions of the n^{th} EP experimental points and the arbitrary point P and

$$H_{ij}=\frac{\partial u_i}{\partial x_j}$$

are the elements of the displacement gradient tensor. It can be decomposed in a symmetric and an anti-symmetric part as $\mathbf{H=E+\Omega}$, where $\mathbf{E}$ is the strain tensor and $\mathbf{\Omega}$ is the rigid body rotation tensor.

A DInSAR interferogram can be related to the unknown components U_i ($i=1..3$) of the displacement vector of the point P according to the following equation:

$$D_{LOS}^{P} = [U_1, U_2, U_3][S_x^P, S_y^P, S_z^P]^T \quad (4)$$

where D_{LOS}^{P} is the LOS displacements, at the point P on the Earth's surface and $[S_x^P S_y^P S_z^P]$ is a unit vector pointing from the point P toward the satellite.
In a compact form the system of equation (1) and (4) can be written as

$$Al = u + e \quad (5)$$

where $l=[U_1\ U_2\ U_3\ \varepsilon_{11}\ \varepsilon_{12}\ \varepsilon_{13}\ \varepsilon_{22}\ \varepsilon_{23}\ \varepsilon_{33}\ \omega_1\ \omega_2\ \omega_3]$ is the vector of unknown parameters, $u = [u_{1(1)} u_{2(1)} u_{3(1)} ... u_{1(n)} u_{2(n)} u_{3(n)} D_{LOS}^{P}]$ is the observation vector, e is the residual vector which model the stochastic nature of the estimation problem and A is the design matrix defined as

$$A = \begin{bmatrix}
1 & 0 & 0 & \Delta x_{1(1)} & \Delta x_{2(1)} & \Delta x_{3(1)} & 0 & 0 & 0 & 0 & \Delta x_{3(1)} & -\Delta x_{2(1)} \\
0 & 1 & 0 & 0 & \Delta x_{1(1)} & 0 & \Delta x_{2(1)} & \Delta x_{3(1)} & 0 & -\Delta x_{3(1)} & 0 & \Delta x_{1(1)} \\
0 & 0 & 1 & 0 & 0 & \Delta x_{1(1)} & 0 & \Delta x_{2(1)} & \Delta x_{3(1)} & \Delta x_{2(1)} & -\Delta x_{1(1)} & 0 \\
. & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . \\
1 & 0 & 0 & \Delta x_{1(N)} & \Delta x_{2(N)} & \Delta x_{3(N)} & 0 & 0 & 0 & 0 & \Delta x_{3(N)} & -\Delta x_{2(N)} \\
0 & 1 & 0 & 0 & \Delta x_{1(N)} & 0 & \Delta x_{2(N)} & \Delta x_{3(N)} & 0 & -\Delta x_{3(N)} & 0 & \Delta x_{1(N)} \\
0 & 0 & 1 & 0 & 0 & \Delta x_{1(N)} & 0 & \Delta x_{2(N)} & \Delta x_{3(N)} & \Delta x_{2(N)} & -\Delta x_{1(N)} & 0 \\
S_x^P & S_y^P & S_z^P & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \quad (6)$$

Assuming a uniform strain field a suitable method to solve the system of equations (5) is the Weighted Least Squares (WLS) which gives the expression (6) as a suitable formula to estimate the unknown vector l

$$\hat{l} = (A^T W A)^{-1} A^T W u \quad (6)$$

where W is the data covariance matrix. Usually W is assumed to be diagonal, i.e. of the form

$$W = diag(\sigma_{1(1)}^{-2}, \sigma_{2(1)}^{-2}, , \sigma_{3(1)}^{-2} ..., \sigma_{1(n)}^{-2}, \sigma_{2(n)}^{-2}, \sigma_{3(n)}^{-2}) \quad (6)$$

where the quantities $\sigma_{j(n)}$'s are the standard deviations of the measurements. According to the modified least squares (MLS) approach proposed by [7], based on the adjustment of the covariance matrix W, we use the matrix **W'** which is a weighted version of the matrix **W** of experimental data. Following the suggestion given by ref. [6, 7] the weight function considered here is:

$$W' = W \exp(-\frac{d_{(n)}}{d_0}) \quad (7)$$

where $d_{(n)}$ is the distance between the n^{th} EP and the arbitrary point P, and d_0 is a distance-decaying constant defining the "level of locality" of the estimation.

Clustering analysis method

The Self-Organizing Map (SOM), also known as Kohonen map, is a popular neural network based on unsupervised learning [8] useful in data visualization an exploration. The SOM maps high-dimensional input vectors onto two-dimensional grid of prototype vectors that are easer to visualize and explore than the original data. The SOM approach is essentially based on the competition between the nodes and it is characterized by a modified version of the winner take all algorithm in which not only the winner node is updated but also its neighbors.

The most commonly used methods to cluster the SOM visually is based on the U-matrix. It visualizes distances between neighbouring map units, and thus show the cluster structure of the map. The highest values of the U-matrix indicate a cluster border while uniform areas of low values indicate clusters themselves [10]. By studying the final U-matrix map, and the underlying features planes of the map, a number of cluster can be identified by K-means algorithm [11, 12]. The best clustering structure, which have been obtained from the K-means algorithm, is selected using Davies-Boulding index [13]. This index uses the within-cluster distance and the between-cluster distance. The Davies-Boulding index is suitable for evaluation of K-means partitioning because it gives low values indicating good clustering results.

Case study

In this section results of the clustering analysis are briefly reported. We have statistically found that the optimal number of clusters in which displacements vectors are grouped is six. In particular we can distinguish at least four main different clusters on the eastern flank of the volcano; these clusters define different sectors showing different kinematics. It is interesting to note the agreement between the cluster and the position of the known main faults of the volcano. In fact, the faults border the cluster, confirming their role in decoupling the entire eastern flank of Mt. Etna in several blocks affected by different deformation. 1) The first one (gray) involves all the central sector of the flank and shows a marked seawards motion with the greatest measured velocities. It is bordered by some of the known faults of the volcano; this motion is in fact confined to the North by the easternmost prolongation of the Pernicana fault towards ESE, as already detected by GPS measurements [14]; westrwards, this sector is bordered by the Ripe della Naca faults and the Valle del Bove area; southwards, the S. Venerina – S. Tecla faults close the gray sector. 2) The second sector (red) surrounds the gray one and seems to be confined by two well known active faults: the Fiumefreddo fault on the northeastern side of the volcano and the Fiandaca one on the southern slope. 3) Here, this fault separates the red area from the green one, that involves the entire upper southern flank of Mt. Etna; this area ends southwards along the Nicolosi – Tremestieri alignment and westwards along the Ragalna fault and its upwards prolongation as defined by SAR investigation [15]. However, the South-western area of the volcano was not covered by GPS measures during the investigated period, so the cluster analysis in this area is conditioned only by InSAR data and then is unreliable. 4) On the uppermost northern side of the volcano, the blue cluster define a high velocity area totally confined to the North by the NE-Rift, Provenzana and Pernicana system, confirming the key role played by these faults in driving the ground deformation in this sector of the volcano [16, 2]. The partitions obtained around the summit area agree with the significant inflation affecting the western and upper flanks in the period 2003-2004 [1].

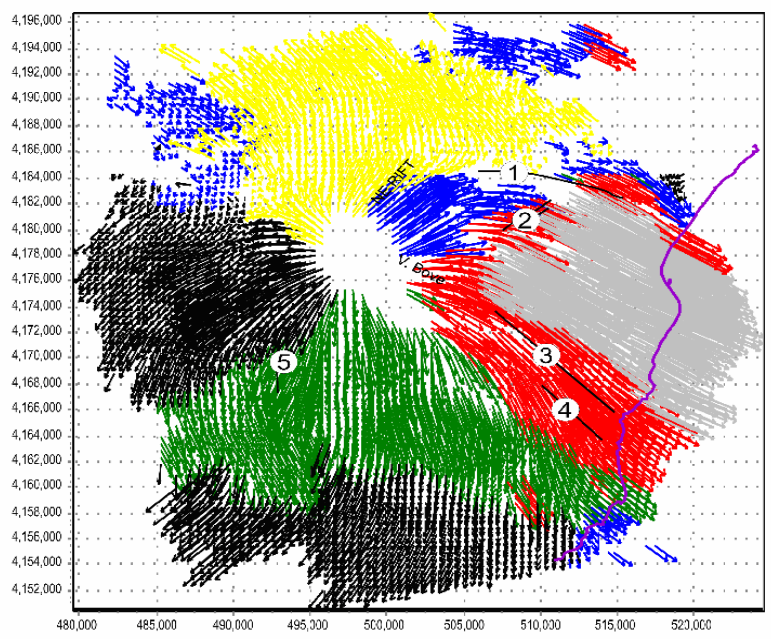

Figure 1. Clustering analysis of vectors displacement estimated by integrating DInSAR data and GPS measurements acquired at Mt Etna in the period 2003-2004. 1): Pernicana fault; RPN; 2) Ripe della Naca faults; 3) S.Venerina- S. Tecla fault; 4) Fiandaca fault; 5)Ragalna fault.

Conclusions

In this paper we have first combined DInSAR data and GPS measurements acquired at Mt Etna in the period 2003-2004 in order to estimate a three dimensional deformation map. Then a clustering analysis based on the combination of SOM and K-means method was applied to the obtained deformation field with the aim to identify sectors of Etna volcano sharing some common deformation patterns. Results show a good agreement between the clusters and the position of the known main faults of the volcano. Furthermore both the seawards motion of the eastern flank of Etna volcano and the significant inflation affecting the western and upper flanks are emphasized by the clustering analysis.

References

[1] Bonaccorso, A., A. Bonforte, F. Guglielmino, M. Palano, and G. Puglisi (2006), Composite ground deformation pattern forerunning the 2004–2005 Mount Etna eruption, J. Geophys. Res., 111, B12207, doi:10.1029/2005JB004206.

[2] Bonforte, A., Gambino, S., Guglielmino, F., Obrizzo, F., Palano, M., Puglisi, G., - Ground deformation modeling of the flank dynamics prior to the 2002 eruption of Mt. Etna. Bull. Volcanol., (2007) vol. 69, pp. 757-768, doi:10.1007/s00445-006-0106-1.

[3] Bonforte, A., Ferretti, A., Prati, C., Puglisi, G., Rocca, F., (2001). Calibration of atmospheric effects on SAR interferograms by GPS and local atmosphere models: first results. J. Atmos. Sol.-Terr. Phys. 63, 1343–1357.

[4] Puglisi, G., Bonforte, A., Guglielmino, F., Palano, M., Prati C., - Dynamics of Mt. Etna before, during and after the July - August 2001 eruption inferred from GPS and DInSAR data. J. Geophys. Res., (2008) vol. 113, B06405, doi:10.1029/2006JB004811.

[5] Guglielmino, F., Nunnari, G., Puglisi, G. and Spata, A., submitted to IEEE Remote sensing and geoscience

[6] Teza, G., Pesci, A., Galgaro, A.,(2008) Grid_strain and grid_strain3: Software packages for strain field computation in 2D and 3D environments, Computers & Geosciences, Volume 34, Issue 9, September 2008, Pages 1142-1153, ISSN 0098-3004, DOI: 10.1016/j.cageo.2007.07.006.
[7] Shen, Z.-K., Jackson, D.D., and B.X. Ge (1996). Crustal deformation across and beyond the Los Angeles basin from geodetic measurements, Journal of Geophysical Research **101**, pp. 27957–27980
[8] Kohonen, T., Self-Organizing Maps, Springer, Berlin, 1995.
[9]Principe, Euliano, and Lefebvre, NeuroSolution, Interactive Book, Neural and Adaptive Systems: Fundamental Through Simulations.
[10] Kohonen, T., Jussi Hynninen, Jari Kangas, Jorma Laaksonen, The Self-Organizing Map Program Package, Helsinki University of Technology, Laboratory of Computer and Information Science, 2005.
[11] Dubes, R., Jain, A.K.,1976. Clustering techniques: the user's dilemma. Pattern Recognition 8, 247-260.
[12] McQueen, J., 1967. Some methods for classification and analysis of multivariate observations. Fifth Berkeley Symposium on mathematics, Statistics and Probability 1, pp. 281-298.
[13] Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Transaction on Pattern Analysis and Machine Intelligence 1(2), 224-227.
[14] Bonforte, A., Puglisi G., 2006. Dynamics of the eastern flank of Mt. Etna volcano (Italy) investigated by a dense GPS network. Journ. Volcanol. Geotherm. Res., 153, 357-369, doi: 10.1016/j.jvolgeores.2005.12.005.
[15] Neri, M., F. Guglielmino, and D. Rust (2007), Flank instability on Mount Etna: Radon, radar interferometry, and geodetic data from the southwestern boundary of the unstable sector, J. Geophys. Res., 112, B04410, doi:10.1029/2006JB004756.
[16] Puglisi, G., Bonforte A., 2004. Dynamics of Mount Etna volcano inferred from static and kinematic GPS measurements. J. Geophys. Res., 109, B11404, doi:10.1029/2003JB002878.