Multivariate approach to evaluate the relationship among geophysical and geochemical variables during an unrest period at Campi Flegrei caldera (Italy)

Abstract
Campi Flegrei (CF) is a caldera in Southern Italy, which has manifested signs of significant unrest in the last years. Indeed, during volcanic crises, the ground of Campi Flegrei caldera begin to grow steadily, the earthquake swarms become more common and strong variations in the chemical composition of fumaroles are observed. In the Campi Flegrei volcanic area there is a spread ground deformation monitoring network. In the years 1983-85 the bradiseism showed an increasing trend of ground uplift with a fairly fast velocity. In the following years (1986-2014) the phenomenon tends to decrease: the earthquakes number is near to zero, the vertical ground displacement decreases, the fumaroles modify the gas emissions, the soil temperature decreases. The main object of this study is to evaluate, by means of variables comparison, in a counterclockwise periods, the relationship among geochemical data of fumaroles, earthquakes and ground deformations, to discover which variables can be considered signals (*geochemical indicator*) of the bradyseismic events during an unrest period.

Key words: geochemical variables, earthquakes, ground deformation, bradyseism, volcanic risk, multivariate analysis

1 Introduction

Magma intrusion in the rocks underneath a volcano generates a number of phenomena, among them, the most relevant are: the increase in the frequency and intensity of micro-earthquakes in the volcano area; soil deformation; compositional changes in fumaroles. Indeed, during volcanic crises, the ground of Campi Flegrei caldera begin to grow steadily, the earthquake swarms become more common and strong variations in the chemical composition of fumaroles are observed (Cardellini et al. 2017; Chiodini. et al. 2011; Chiodini 2010).

The Campi Flegrei (CF) caldera (Fig. 1) represents a suitable subject being a densely populated area where the volcanic hazard is very high. During its history, the caldera has experienced very large explosive events. The eruptions of the Campanian Ignimbrite and of the Neapolitan Yellow Tuff (Deino et al. 2004; Isaia et al. 2009), led to the formation of its primary structure. The more recent Agnano Monte Spina and Astroni explosive eruptions, modified the caldera structure as well as the numerous successive eruptions, which generated pyroclastic deposits spread over an extremely large area (Isaia et al. 2004).

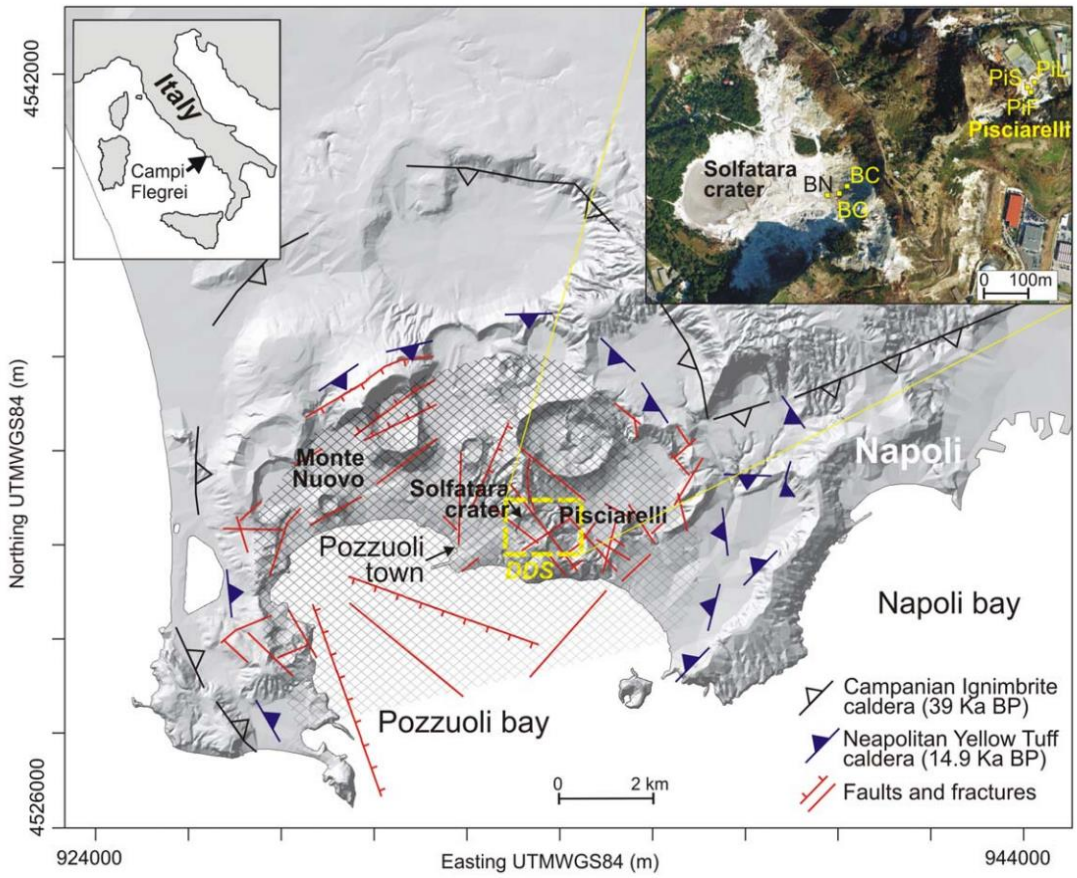

Fig. 1 Campi Flegrei caldera geological map showing the Solfatara crater where the geochemical samples were acquired. The GPS station (RITE) is located near the Pozzuoli town. The reference grid area indicates where the earthquakes in the period 1983-2016 were detected.

Today, multivariate statistical analyses to study volcanic phenomena are more addressed to sectorial data (Hernández-Antonio et al. 2015; Steinhorst et al. 2001; Join et al. 1997; Anazawa et al. 1994; Velasco-Tapia 2014; Jamieson et al. 2015; Buccianti et al. 2015). A multivariate statistical approach applied to interdisciplinary datasets is only present in some works (Kumamoto, et al. 2016, Avery et al. 2017), where it is shown how useful this methods can be for the seismic and volcanic risk assessment. Multidimensional analysis allows, in fact, to explore a subspace of the original variables space where independent phenomena can be more easily identified.

A typical multivariate data set is represented by the collection of geophysical and geochemical observations from the volcanoes monitoring (GPS, INSAR, tiltmetry. water table level, temperature, fluids flux etc.). The attempt is to study the volcanic state

evolution, which has inherent difficulties based both on the physics complexity of the phenomenon and on the strong structure dissimilarity of the volcanoes.

Our basic observation is the recognition of two distinct periods in the ground deformations time series, 1983-1985 and 1986-1988 (Fig. 2), with a reversal trend and a clear turning point centered around 1985. These two periods are used as reference for our statistical comparative analysis, to identify significant differences in the geochemical and geophysical variables values. Our main objective is to discover the latent relationships among all the variables, in order to introduce a complex geo-indicator useful for predicting the evolution of the bradyseismic phenomenon during an unrest period.

The paper is organized as follows. In the first part (Section 2 and 3), we present the used datasets and we apply some statistical tests to model the bradyseismic phenomenon. In the second part (Section 4) we introduce a wavelet analysis as an indispensable tool to strengthen our tests and to analyze datasets in the frequency domain, as well.

In particular, in the first part (Section 2), we introduce the three statistical analyzed datasets: ground deformation, gaseous emissions from fumaroles and earthquakes, recorded in the periods under investigation. As the dataset of the fumaroles gaseous components is a compositional dataset, a logarithmic transformation is chosen, and the Aitchison geometry is applied (Aitchison 1981; 1982) to analyze it. After the transformation of the geochemical variables, a comparative study is made between the two symmetric periods (Section 3), to identify which variables are actively involved in the evolution of the bradyseismic phenomenon. Moreover, the geochemical and geophysical datasets are jointly analyzed in order to highlight the relationships and the latent variables underlying the bradyseismic phenomenon.

To refine the research of the geochemical indicators, the wavelet analysis (Section 4) was considered appropriate, because the analyzed datasets are composed of historical not stationary time series that have to be investigated in the frequency domain. In this work, we suggest the wavelet analysis because ground displacement, geochemical variables, and earthquakes are non-stationary time series. The wavelet analysis have never been applied before to geochemical and geophysical variables in this volcanic area.

Finally, in the last Section, there are the conclusions.

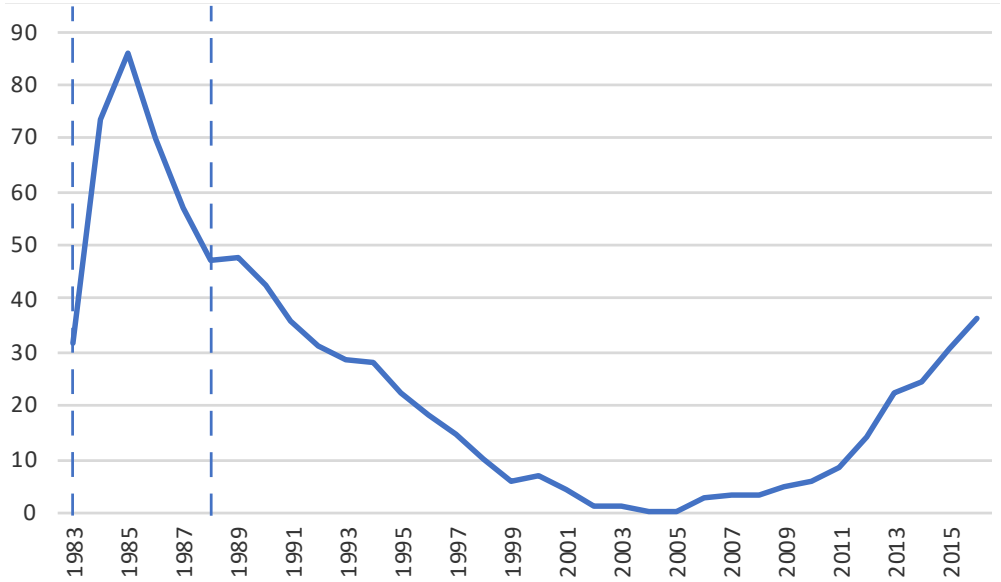

Fig. 2 Ground displacement (cm), from 1983 to 2016 at RITE GPS station. The process of increase and decrease of the ground between 1983 and 1988 (dotted line at left side) was selected to study the unrest as function of geochemical and geophysical variables

2 Datasets

The variables selection was done following three basic principles: 1) relevance respect to the problem; 2) exhaustiveness in the phenomenon representation; 3) not redundancy. The data of ground uplift and downlift represent the first dataset that is a reference for the geochemical variables (second dataset) and the earthquakes (third dataset) in the area of bradyseims (Fig. 1).

In the Campi Flegrei volcanic area there is a spread ground deformation monitoring network. In our work, we refer to a single GPS detection station, located in Pozzuoli, Rione Terra (RITE, Figure 1), as it is representative of the whole deformation trend. The acquisitions of the ground deformation data from the RITE station between 1983 to 2016 is shown in Figure 2 on the left side, there we can see the last remarkable event occurred between 1983 and 1985.

The monthly average of vertical ground deformation time series, represents our *first dataset* (see Table A.1). From 1983 to 2016, the maximum annual average displacement (85 cm) was achieved in 1985. In this paper, we evaluate and compare the ground uplift period (1983-1985) with the following period of ground downlift (1986-1988) (Fig. 2).

This study uses, as *second dataset*, the time series of geochemical variables collected in Solfatara volcano from Bocca Grande fumarole between 1983-1988 (Tables A.1 and A.2). The gas detecting system (geochemical variables), constantly used by the

monitoring staff of the National Geophysical and Vulcanology Institute, the Vesuvian Observatory Section, envisages the capture of the fumes of gases emitted by the fumarole in a single special ampoule. The capture is suspended after a predetermined time interval (seconds) and this time is the same for all samplings. The gases taken in the ampoule are then evaluated in micromole / mole, but being part of a whole (single sample) they are compositional data (Aitchison, 1986). From each sample we obtain the percentages of gas (H_2O, CO_2, N_2, H_2S, H_2, CH_4) that correspond to a gas concentration. The variables are non-uniformly sampled and organized in monthly averaged series to be analyzed in a multivariate scheme. The sum of theirs percentages is a constant k for all samples (k = *one million*). The geochemical variables can be collected in a matrix X of n rows and p columns: each column represents a geochemical variable X_j and each row X_i contains the sample of the observations x_{ij} *(i=1,2,…,n; j=1,2,…,p)* at a given date. These geochemical variables, indicative of the bradiseismic phenomenon evolution, are compositional data since form a "closed set" of p dimensions and assume meaning only if they are evaluated within the closure itself (Filzmoser et al. 2009). Specific geometry is required for analyzing such compositional data (Aitchison 1981; 1982). The Aitchison metric, is an *isometric transformation* of the Simplex in a subspace of p dimensions with Euclidean metric (Egozcue et al. 2003). In this paper, we transform the geochemical variables by means of the *centred log-ratio (clr*) that is the transformation of X in:

$$\boldsymbol{X} = \left[\ln(x_{ij}/g_i)\right] (i = 1,2,\dots,n; j = 1,2,\dots,p) \quad (1)$$

where g_i is the geometric mean of each composition (row of X). In Appendix are shown the transformed dataset by means of *clr* transformation. In the following, each geochemical component, transformed using clr, will be cited using the suffix "clr".

Earthquakes recorded in the Campi Flegrei caldera area (fig. 1) between 1983 and 1988 form the *third dataset*. In the entire considered period there is an inhomogeneous density distribution of earthquakes: between 1983 and 1984 several earthquake events were recorded there while no seismic event were in the second period (1986-1988) (see Appendix). Between 1983-1985, 7687 VT seismic events were located in the Pozzuoli area (Fig. 1) at depths from 0 to 4 km. Gutenberg-Richter distribution (Gutenberg and Richter 1954) (Fig. 3) well fit to the data with magnitude (M) M>=0.5. In this study we map the minimum of the frequency magnitude distribution, for the catalog completeness, at the 93% level (R^2 = 0.96). We select the M at which 93% of the observed data (5924 events) are modeled by a straight line fit (Wiemer and Wyss 2000). The straight line in Fig. 3 represents the theoretical Gutenberg-Richter distribution.

The transformation of geochemical variables is fundamental to jointly study all the variables avoiding the closure existing in the geochemical dataset. The basic hypothesis, in fact, is that geochemical variations and earthquakes are correlated to the ground deformation. All the three datasets are divided in two subgroups that cover respectively, years 1983 to 1985 and years 1986 to 1988 (Tables A.1 and A.2).

In the following section, therefore, we try to identify which variables are actively involved in the evolution of the bradyseismic phenomenon by comparing the values of the first and the second period in which we have detected the inversion of the soil deformation trends.

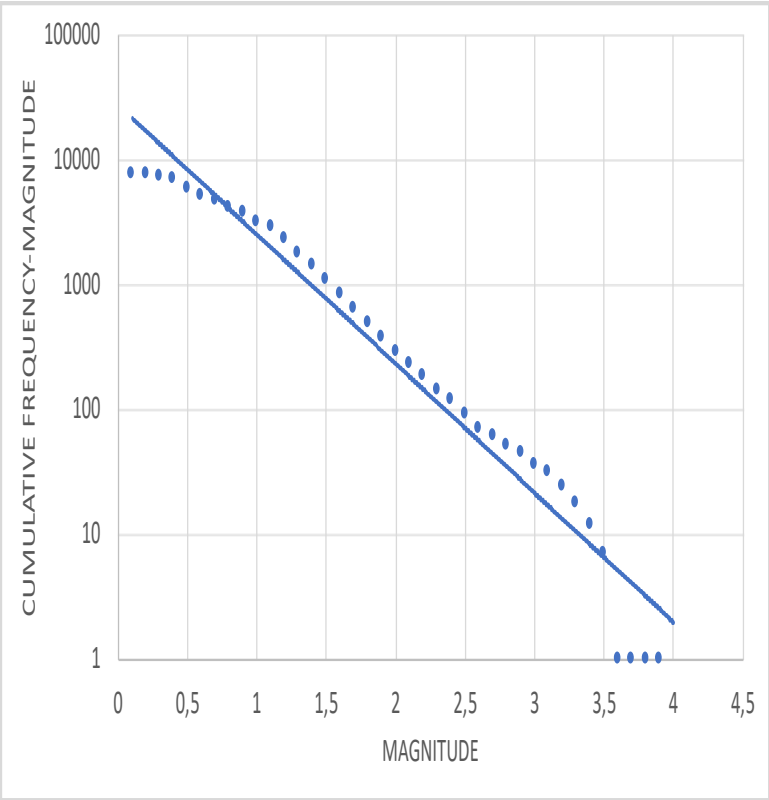

Fig 3 Gutenberg-Richter distribution for the CFc catalog. The x-axis is the magnitude while the y-axis represents the log of cumulative frequency-magnitude distribution of earthquakes (1983-1995). The straight line represents the theoretical Gutenberg-Richter distribution.

3 Analysis of geochemical and geophysical variables in two periods

The *centred log-ratio (clr*) was applied to the geochemical dataset to obtain the Centers for the first and the second subgroups, respectively $Cen(X_I)$ and $Cen(X_{II})$. The centers are measures similar to the averages, as proposed by the Aitchison geometry (1981; 1982). In Tab. 1 are shown and compared the Centers between the I*st* and II*nd* period together with the two-sample Kolmogorov-Smirnov (K-S) test for equality of distribution function and the two-sample Wilcoxon rank-sum (Mann-Whitney) test.

H_2S and H_2O are the main geochemical components, which have a significant high variation between the unrest period and the successive period (Tab. 1); we note that H_2S shows a sharp decrease in the quantity emitted by the fumaroles, reducing its amount of 33.58%. The H_2O and CO_2 reduction is more moderate, respectively 17% and 8.7%. Vice versa, in the period of decreasing ground uplift (since 1986) the other gaseous elements CH_4, N_2 and H_2 reemerge significantly with quantities respectively of 32.31%, 6.66% and 2.42%.

As a further variable, the *ground displacement velocity* (GDV) is introduced since it should provide useful indications on the inference of the bradyseismic phenomenon. Table 2 shows the correlations between geochemical variables, ground displacement (GD), earthquakes (EQ) and ground displacement velocity, in the first period of the bradyseismic unrest (1983-85).

The comparative analysis of the correlations between all the geochemical and geophysical variables, in the first and second period (Tab. 2) leads to interesting results. CO_2 (r = .92; .41) and N_2 (r = .85; .93) are positively and highly correlated with the **ground displacement (GD)**, respectively, in the first and in the second period. Viceversa, H_2 (r = -.87; -.65) and H_2S (r = -.54; -.89) have a negative correlation in the two periods. CH_4 and H_2O have a different behavior with respect to the other geochemical components, as CH_4 is negatively correlated with the ground uplift in the first period of reference (r = - 84), while H_2O is negatively correlated in the second period (r = -.71). An increase in CO_2 and N_2 is associated with the increase in temperature measured in volcanic areas (Chiodini et. al. 2010; 2011), as well as the ground uplift.

Worth of note are the correlations found among the **ground displacement velocity (GDV)**, H_2S (r = .89; .85) and H_2O (r = .78; .73), positively correlated, both in the first and in the second period, while H_2 is only correlated (r =.73) with the displacement velocity in the second period. In particular, the high correlation between H_2S, the ground displacement velocity (r = .89) and a decrease in CO_2 (r= -.49) denote an excellent signal of an increased volcanic activity. A negative correlation with ground displacement is detected for N_2 (r = -.79;-.73) in both the reference periods while for CO_2 only in the first period.

As concerned the correlations with **earthquakes,** in the first period, H_2S is still positively correlated (r = 0.88) as well as the ground displacement velocity (r = 0.87) and H_2O (r = 0.67). Negatively correlated are N_2 (r = -0.85), CO_2 (r= -.58) and the ground displacement (GD) (r=-.56). In the second period of reference, since the earthquakes with a magnitude greater than 0.5 were completely absent, the correlations could not be calculated and the symbol "()" in Tab. 3 indicates this absence. The correlations between earthquakes, H_2S and the ground displacement velocity leads to underlying the strong association between these components.

Table 1: Centers between I*st* and II*nd* period with two-sample Kolmogorov-Smirnov (K-S) test for equality of distribution function and two-sample Wilcoxon rank-sum (Mann-Whitney) test

	H_2O clr	CO_2 clr	H_2S clr	N_2 clr	CH_4 clr	H_2 clr
Cen(X_I)	5.73	4.23	-0.66	-1.52	-5.15	-2.62
Cen(X_{II})	5.55	4.14	-1,00	-1.32	-4.83	-2.56
(Cen(X_{II})- Cen(X_I))*100	-17.11	-8.70	-33.58	2.42	32.31	6.66
K-S Cen(X_I)	0.03	0.25	0.00	0.45	0.61	0.61
p-value	0.97	0.10	1.00	0.00	0.00	0.00
K-S Cen(X_{II})	-0.72	-0.59	-0.71	-0.40	-0.03	-0.05
p-value	0.00	0.00	0.00	0.00	0.97	0.93
Combined K-S	0.72	0.59	0.71	0.45	0.61	0.61
p-value	0.00	0.00	0.00	0.00	0.00	0.00
Wilcoxon rank-sum: P[Cen(X_I)>Cen(X_{II})]	0.92	0.65	0.92	0.50	0.22	0.30

Furthermore, as concerned the significant differences between the I.st and the II.nd period, only the correlations between CH_4 and H_2 decrease from .87 to -.78 showing an inversion of the significant correlation in both the reference periods.

In order to obtain a more immediate and complete picture of the differences in the geochemical components behavior with respect to the geophysical variables, an analysis has been made in Principal components for the whole reference period (1983-88). A Biplot (Aitchinson and Greenacre 2002) was performed for geochemical (*clr* transformed), earthquake, ground displacement and ground displacement velocity (all standardized variables) variables. Looking at the biplot (Fig. 4) of the first period (1983-85) the differences in the signals are evident: the increase in H_2S and H_2O and earthquakes are related to the *bradyseism velocity*, while CO_2 is related to the vertical *ground displacement speed* in the second period, CH_4 is in relation with H_2. The total explained variance in the Biplot is 0.90.

To investigate more in details on these geochemical and geophysical variables and their correlations, we have considered to perform a wavelet analysis never applied before in this volcanic area.

Table 2: Comparison of correlations among Geochemical, Ground Displacement (GD), Ground Displacement Velocity (GDV) and Earthquakes (EQ) from 1983 to 1985 and from 1986 to 1988 (in parenthesis). With * is shown the correlation significant at level 0.05 (2-Queue). The symbol () indicate the impossibility of calculating the correlations due to the absence of earthquakes

	H_2Oclr	CO_2clr	H_2S clr	N_2 clr	CH_4 clr	H_2 clr	GD	GDV	EQ
H_2Oclr	1.00								
CO_2clr	.03 (.25)	1.00							
H_2Sclr	.82* (.87*)	-.47* (-.12)	1.00						
N_2clr	-.48 (-.55*)	.83* (.50*)	-.87* (-.75*)	1.00					
CH_4clr	-.38 (-.70*)	-.92* (-.72*)	.13 (-.46*)	-.59* (-.16)	1.00				
H_2clr	-.29 (.86*)	-.87* (.23)	.26 (.82*)	-.65* (-.40*)	.87* (-.78*)	1.00			
GD	-.02 (-.71*)	.92* (.41*)	-.54* (-.89*)	.85* (.93*)	-.84* (.10)	-.87* (-.65*)	1.00		
GDV	.78* (.73*)	-.49* (-.32)	.89* (.85*)	-.79* (-.73*)	.11 (-.28)	.17 (.73*)	-.43* (-.88*)	1.00	
EQ	.67* ()	-.58* ()	.88* ()	-.85* ()	.28 ()	.37* ()	-.56* ()	.87* ()	1.00

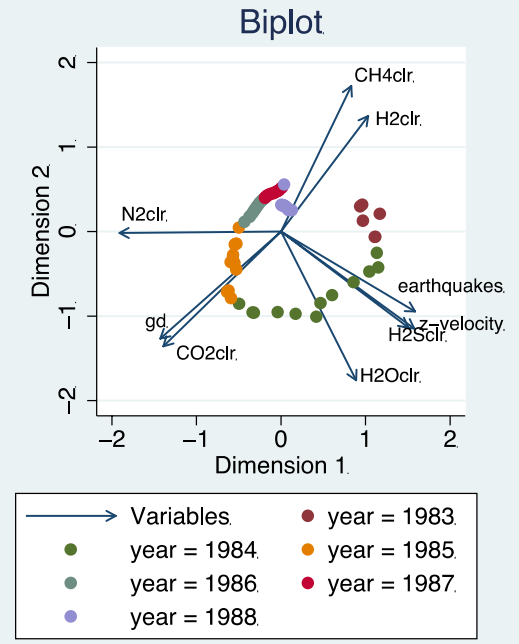

Fig. 4 Biplot of geochemical, earthquakes, ground displacement (GD) and ground displacement velocity (z-velocity, GDV) variables (1983-1988).

4 Wavelet Analysis

Many times, the information that cannot be readily seen in the time-domain can be accessed in the frequency domain. Fourier transform (FT) is used in many different areas including all branches of engineering. Hilbert transform, short-time Fourier transform, Wigner distributions, the Radon Transform, and the Wavelet transform (WT) are available.

The wavelet analysis is a very powerful tool to explore nonstationary phenomenon (Daubechies 1990); it provides a map of the signal decomposed in a time-scale space. The 2D representation of the signal allows for a quick identification of nonstationary part of the signal and relative scale as well as for the stationary part.

Let **x(t)** be the signal to be analyzed, this can be expanded in a wavelet base. The wavelet base element, to be "admissible" as a wavelet, must have zero mean and be localized in both time and frequency space (Farge 1992).

The definition of Continuous Wavelet Transform *C* is:

$$C_x^{\varphi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t)\, \varphi^*[(t-\tau)/s]dt \tag{2}$$

where $\varphi^*[(t-\tau)/s]$ is called the mother function wavelet and the (*) indicate the complex conjugate (note that the mother function is invariant for translation). The mother function wavelet is a prototype for generating the other window functions: all the used base elements are expansion, compression and translation of the mother function; $\boldsymbol{\tau}$ is the translation parameter (localize the wavelet on the variable axes), *s* the scale parameter (measure of period: *s* >1 dilate the mother wave and *s* < 1 compress the mother wave); $1/\sqrt{|s|}$ is a normalization constant.

Also assume that a wavelet function depends on a nondimensional "time" parameter. The CWT is a measure of similarity between the basis functions (wavelets) and the signal itself.

The procedure will be started from scale **s=1** and will continue for the increasing values of **s** , i.e., the analysis will start from short period (high frequencies) and proceed towards long period (low frequencies). This first value of **s** will correspond to the most compressed wavelet. Wavelet analysis produces a time-scale view of the input signal. The location of the wavelet allows to explicitly represent the location of events in time. The shape of the wavelet allows to represent different detail or resolution.

In our study, we first have analyzed the correlation structure with Pearson coefficients, in order to obtain the global correlation amongst the variables. Then we have gone into details of this structure analyzing, by the wavelet, the inner patterns to extract local simultaneous transients as well as common oscillatory signals. The wavelet analysis involves a larger period ranging from 1983 to 2016 to obtain an overall picture of the phenomenon under study. Our analysis uses the correlation coefficients $R_W^2(s,\tau)$ and phases $\theta_W(s,\tau)$ between a couple of variable (x_1, x_2) as defined by Torrence and Webster 1999);

$$R_W^2(s,\tau) = \frac{|\langle s^{-1} W^{x_1 x_2}(s,\tau)\rangle|^2}{\langle s^{-1}|W^{x_1}(s,\tau)|^2\rangle\langle s^{-1}|W^{x_2}(s,\tau)|^2\rangle} \quad (3)$$

$$\theta_W(s,\tau) = tan^{-1}\frac{Im(s^{-1}\langle W^{x_1 x_2}(s,\tau)\rangle)}{Re(s^{-1}\langle W^{x_1 x_2}(s,\tau)\rangle)} \quad (4)$$

where:

$$W^{x_1 x_2}(s,\tau) = C^{x_1}(s,\tau)C^{x_2 *}(s,\tau) \quad (5)$$

represents the cross power between the wavelet coefficients of the two series defined in eq. 2) and $\langle\ \rangle$ denotes a smoothing operator in time and scale. The graphic representation of the coherence between couple of variables has made via a color scale, which indicate the amplitude of it. The relative phases have been represented via black arrows, the convention is such that if the two variables, at a generic fixed point of the graph, are in phase (lag = zero) the arrow is horizontal pointing toward positive x-axes. If the first variable precedes the second the arrow rotates clockwise (counterclockwise in opposite case). Variables in phase opposition (180 degrees) are represented by horizontal arrow pointing toward negative x-axes. The white dashed line delimits the area (COI, cone of influence) where there is not bias, dues to edge effects, on the estimated coherence. At high period (high part of the graph) good estimates of coherences and phases are restricted. In the figures from 5 to 11, the ordinate axis represents the period relative to the Fourier transform peak frequency of the considered wavelet.

Figures 5 and 6, respectively, show the relationship between earthquakes with ground displacement and earthquakes with ground velocity. It is clear the marked value of correlation between the two groups of two variables in both graphs at long period and up to about the 2005. Note the delay of the number of earthquakes with respect to the ground displacement (fig.6, arrows in the yellow zone downward directed) and with a small advance in case of earthquake/Ground displacement couple (fig.7, arrows in the yellow zone horizontal or slightly projected onto positive y-axes).

Both wavelet analysis confirm the correlation (Tab. 2) between earthquakes and ground lifting (r = -0.56) against the positive correlation between earthquakes and lifting velocity (r = 0.87). The structure of the coherent signal at lower periods is more articulate and shows nuclei of aggregation in a different scale/time diagram zones, with, in any case, sustained coherence structure. These zones are localized at the same scale-time interval in the two graphs and are separated, from the longer period high coherence zones, by a stripe of very low coherence (the about 30 degrees blue stripe in both graphs).

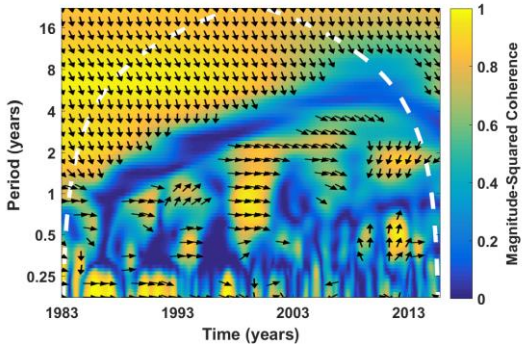

Fig. 5 Ground Displacement and Eartquakes

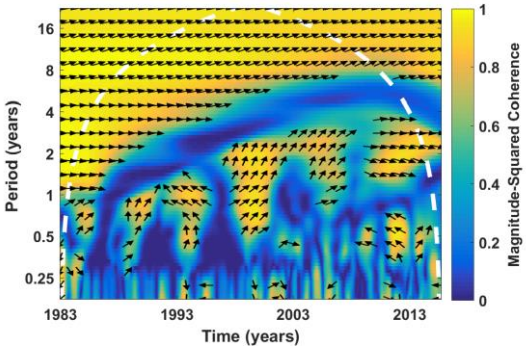

Fig. 6 Velocity of Ground Displacement and Earthquakes

In the following 7 and 8 figures, we show the correlation structure between H_2S and respectively ground displacement (GD) and ground displacement velocity (GDV). The negative correlation (r=-.54) between H_2S and the ground displacement (Fig.8) is reflected in the coherence at very long period (8 ÷ 16 years) and even in the horizontal yellow stripe at period of about 4 years until 2003. The same considerations hold for the Fig.9 with similar structure at high period (r = .89 is the Pearson correlation coefficient up to 1989, Tab. 2). The marked differences between the two comparisons are, at period from 2 to 16 years, substantially in the phases which are in Fig. 8, around 90 degrees (delayed processes), while in Fig. 9 around 0 degrees (process in phases). In any case what have to be noted is that the higher correlation values are reached around the beginning of the Campi Flegrei caldera deformation crises. In the period between the years 2000 ÷2016 the occurrence of high correlation is more spurious and time limited showing isolated period-time spots.

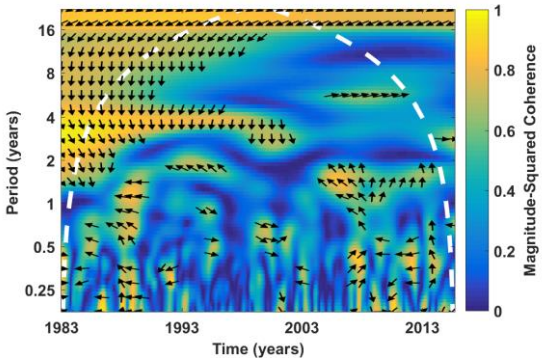

Fig. 7 Ground Displacement and H_2S

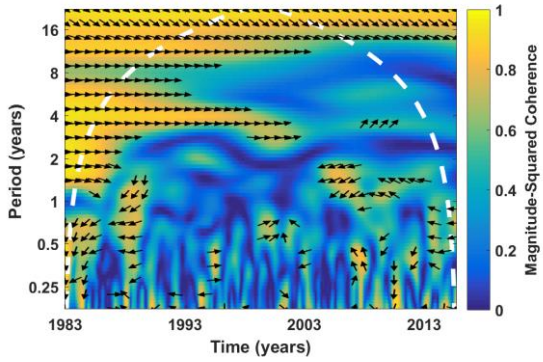

Fig. 8 Ground Velocity and H_2S

The wavelet analysis showed in Fig.9 confirms the trend, perfectly in phase, in the period 1983-1986 as regard the scales (period) 2 ÷ 7 years, and then, lack of correlation due to the lack of earthquakes in the following period. For this reason, H_2S could be

designated as an excellent indicator both of earthquakes and of increase in ground speed and as an indicator of bradyseismic disorders.

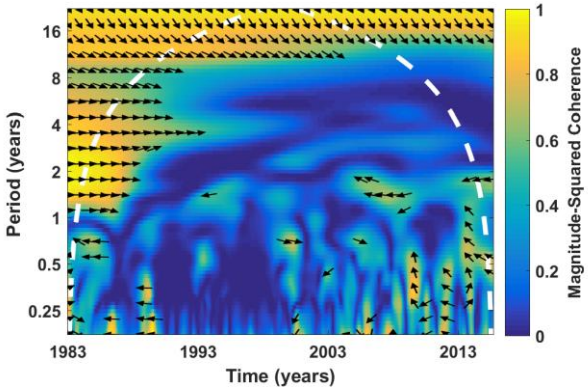

Fig. 9 Earthquakes and H_2S

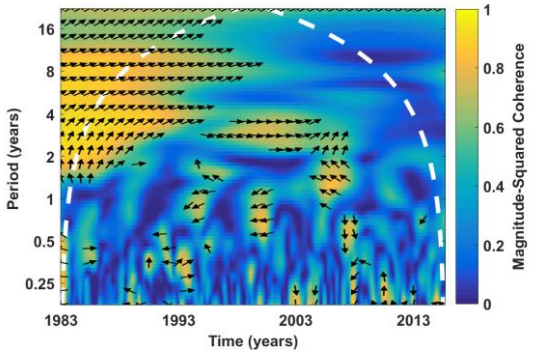

Fig. 10 Ground Displacement and N_2

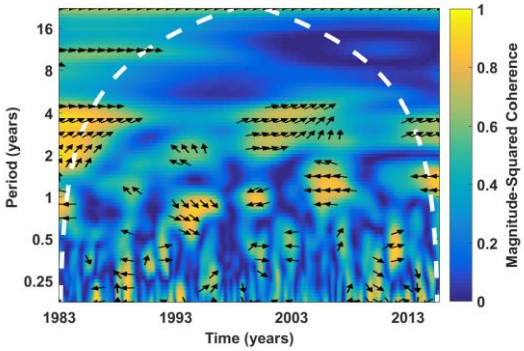

Fig. 11 Ground Displacement and CO_2

Finally, we show in figure 10 and 11 the correlation between the ground displacement and the geochemical species N_2 and CO_2. The coherence is again strong at the beginning of the analyzed period, as regard the N_2, while CO_2 has a general weaker coherence. The graphs give a picture in any case coherent with those showed in Fig. 8 and Fig. 9.

Conclusions

In the first part of the years 1983-85, the bradyseismic phenomenon showed an increasing trend with the ground starting to rise at a fairly fast velocity. After 1985, and in the following years, the bradyseismic trend tends to decrease. Indeed, in the first period the ground displacement increases positively up to 1985, when we have a turning point, with a decrease of the ground level, which reached the previous level.

The choice of the reference periods, characterized by symmetric opposite trends, allowed us to compare the values of the geochemical data, after the logarithmic transformation of the geochemical variables, because they belong to the same subspace.

The main purpose of this study is to evaluate, by means of the comparison between geochemical and geophysical datasets, along the opposite trends of the ground displacement, the relationships among all these variables to discover one or more indicators useful to monitor the evolution of the bradyseismic phenomenon.

In the first part of this work, since the measurements of geochemical variables are samples taken over time, and each sample has a constant sum, they are time series of compositional data, so an appropriate logarithmic transformation (*clr*) was applied for the adoption of the usual statistical techniques.

The *first conclusion* derives by a comparative analysis of the centers. We can affirm that among the geochemical variables, H_2S and H_2O are the main geochemical components, which have a significant high variation in the first period of unrest respect to the successive period, followed by CO_2. We note that H_2S shows a sharp decrease in the amount emitted by the fumaroles, reducing its emission by 34%. In the second period, (soil uplift, since 1986) the other gaseous components CH_4, N_2 and H_2 reemerge significantly (Table 1).

The subsequent analysis of the correlations between geochemical variables with respect to the geophysical variables (Tab. 2) (ground displacement, ground displacement velocity and earthquakes) leads to a *second interesting conclusion*: we note a differentiated behavior of the correlation among the geochemical variables, the displacement of the ground level and the ground uplift speed. Considering that, the velocity is more decisive in assessing *in peius* the bradyseismic system, the correlations with velocity of the ground displacement have been given more importance. Indeed, CO_2 and N_2 are positively and highly correlated with the ground displacement, both in the first and in the second period, while H_2S and H_2O are positively and highly correlated with the velocity of ground displacement.

Regarding the correlations with earthquakes, in the first period of reference, H_2S, H_2O and the ground displacement velocity are positively correlated with the number of earthquakes with magnitude greater that 0.5, which means that they are in continuous association with the ground level variations and that, therefore, their variations can be considered indicative as a bradyseismic signal.

The PCA analysis, visualized by a biplot, leads to a better identification of the relationships and of the subgroups of geochemical elements, which characterize the two periods. In the first period H_2O, earthquakes and ground displacement velocity dominate the scenario; in the second period CO_2 and N_2 are related to the ground displacement, while, the sub-group formed by CH_4 and H_2 is present in every period without being directly linked to the variations in lifting, velocity and earthquakes.

The wavelet analysis have been conducted to confirm the correlation analysis performed in the years from 1983 to 1988 and to obtain an overall picture of the phenomenon under study.

Concluding, there are significant relationship among geochemical variables, earthquakes, ground uplift and ground displacement velocity. The suggested approach and the showed results are the basis to best interpret and analyze the bradyseismic phenomenon.

Appendix

Table A.1: Geochemical dataset, clr transformed, number of earthquakes and ground vertical displacement at RITE station (1983-1985 years)

Year/Month	H_2Oclr	CO_2clr	H_2Sclr	N_2clr	CH_4clr	H_2clr	Ground displacement (cm)	Earthquakes number
83/6	5,74	4.06	-0.49	-2.19	-4.73	-2.40	16.96	244
83/7	5.71	4.04	-0.47	-2.16	-4.67	-2.44	18.89	188
83/8	5.70	3.99	-0.49	-2.11	-4.61	-2.47	20.83	257
83/9	5.68	4.01	-0.46	-2.14	-4.68	-2.41	22.77	616
83/10	5.69	4.04	-0.43	-2.11	-4.69	-2.51	24.70	397
83/11	5.71	4.02	-0.37	-2.15	-4.69	-2.52	31.52	319
83/12	5.74	3.98	-0.45	-2.08	-4.68	-2.51	38.33	381
84/1	5.77	4.03	-0.40	-2.11	-4.77	-2.51	45.15	505
84/2	5.80	4.04	-0.40	-2.07	-4.84	-2.53	51.96	658
84/3	5.82	4.06	-0.39	-2.07	-4.87	-2.55	58.11	570
84/4	5.82	4.13	-0.36	-2.01	-4.99	-2.59	62.93	537
84/5	5.86	4.19	-0.32	-1.89	-5.21	-2.63	67.76	272
84/6	5.87	4.23	-0.36	-1.78	-5.31	-2.65	72.58	243
84/7	5.88	4.27	-0.40	-1.68	-5.41	-2.66	77.40	389
84/8	5.84	4.29	-0.55	-1.44	-5.42	-2.71	82.22	316
84/9	5.81	4.31	-0.70	-1.25	-5.43	-2.74	87.04	269
84/10	5.79	4.34	-0.75	-1.11	-5.45	-2.82	90.75	117
84/11	5.79	4.36	-0.76	-1.10	-5.50	-2.80	91.91	284
84/12	5.79	4.38	-0.84	-1.07	-5.48	-2.78	93.07	58
85/1	5.80	4.40	-0.93	-1.04	-5.47	-2.76	93.48	9
85/2	5.76	4.41	-0.86	-1.04	-5.53	-2.74	91.98	5
85/3	5.72	4.42	-0.79	-1.04	-5.58	-2.72	90.56	8
85/4	5.67	4.36	-0.85	-1.13	-5.36	-2.68	89.26	4
85/5	5.64	4.36	-0.90	-1.10	-5.34	-2.66	87.96	0
85/6	5.62	4.36	-0.95	-1.06	-5.32	-2.64	86.66	0
85/7	5.64	4.34	-0.95	-1.01	-5.33	-2.69	85.25	0
85/8	5.61	4.33	-0.93	-1.02	-5.36	-2.64	83.62	0
85/9	5.58	4.32	-0.95	-1.05	-5.29	-2.60	81.99	0
85/10	5.57	4.32	-0.96	-1.05	-5.24	-2.64	80.35	0
85/11	5.57	4.32	-0.97	-1.05	-5.20	-2.67	78.85	0
85/12	5.50	4.31	-1.01	-1.07	-5.14	-2.58	77.66	0

Table A.2: Geochemical dataset, clr transformed, number of earthquakes and ground vertical displacement at RITE station (1986-1988 years)

Year/Month	H_2Oclr	CO_2clr	H_2Sclr	N_2clr	CH_4clr	H_2clr	Ground Displacement (cm)	Earthquakes
86/1	5.53	4.25	-1.05	-1.08	-5.08	-2.57	76.5	0
86/2	5.56	4.18	-1.08	-1.08	-5.02	-2.56	75.3	0
86/3	5.52	4.21	-1.02	-1.13	-5.00	-2.58	74.1	0
86/4	5.53	4.18	-1.05	-1.14	-4.95	-2.58	72.9	0
86/5	5.53	4.18	-1.05	-1.16	-4.91	-2.58	71.7	0
86/6	5.52	4.17	-1.05	-1.19	-4.87	-2.59	70.5	0
86/7	5.52	4.17	-1.05	-1.21	-4.84	-2.59	69.3	0
86/8	5.52	4.16	-1.05	-1.23	-4.8	-2.59	68.1	0
86/9	5.51	4.16	-1.05	-1.26	-4.77	-2.59	66.9	0
86/10	5.51	4.15	-1.05	-1.28	-4.74	-2.60	65.8	0
86/11	5.51	4.15	-1.05	-1.3	-4.71	-2.60	64.7	0
86/12	5.51	4.15	-1.05	-1.31	-4.7	-2.60	63.6	0
87/1	5.51	4.14	-1.05	-1.31	-4.69	-2.60	62.5	0
87/2	5.51	4.14	-1.05	-1.31	-4.69	-2.60	61.4	0
87/3	5.51	4.14	-1.05	-1.32	-4.68	-2.60	60.4	0
87/4	5.51	4.13	-1.05	-1.32	-4.68	-2.60	59.3	0
87/5	5.51	4.12	-1.04	-1.33	-4.67	-2.60	58.2	0
87/6	5.52	4.10	-1.03	-1.33	-4.67	-2.60	57.2	0
87/7	5.52	4.09	-1.02	-1.33	-4.66	-2.60	56.4	0
87/8	5.52	4.08	-1.01	-1.34	-4.66	-2.60	55.6	0
87/9	5.53	4.06	-1.00	-1.34	-4.65	-2.60	54.7	0
87/10	5.53	4.05	-0.99	-1.35	-4.65	-2.60	53.9	0
87/11	5.52	4.05	-0.97	-1.36	-4.66	-2.58	53.1	0
87/12	5.51	4.06	-0.96	-1.38	-4.68	-2.56	52.2	0
88/1	5.52	4.05	-0.97	-1.39	-4.69	-2.52	51.4	0
88/2	5.6	4.12	-0.91	-1.47	-4.75	-2.60	50.6	0
88/3	5.61	4.14	-0.96	-1.45	-4.78	-2.57	49.7	0
88/4	5.62	4.14	-0.95	-1.44	-4.82	-2.55	48.9	0
88/5	5.62	4.15	-0.94	-1.44	-4.86	-2.53	48.1	0
88/6	5.63	4.15	-0.93	-1.43	-4.91	-2.51	47.3	0
88/7	5.63	4.15	-0.92	-1.43	-4.95	-2.48	46.5	0
88/8	5.63	4.16	-0.91	-1.42	-5.00	-2.46	45.8	0
88/9	5.64	4.16	-0.9	-1.41	-5.06	-2.44	45.1	0
88/10	5.65	4.17	-0.88	-1.4	-5.11	-2.42	44.3	0
88/11	5.65	4.18	-0.87	-1.39	-5.17	-2.40	43.6	0
88/12	5.67	4.16	-0.89	-1.38	-5.19	-2.38	43,5	0

References

Aitchison, J.: A new approach to null correlations of proportions. Mathematical Geology. 13, 175-189 (1981)

Aitchison, J.: The statistical analysis of compositional data (with discussion). Journal Royal Statistical Society. B44, 139-177 (1982)

Aitchinson, J.: The Statistical Analysis of Compositional Data. London: Chapman and Hall (1986)

Aitchinson, J., Greenacre, M.: Biplots of compositional data. Journal of the Royal Statistical Society: Series C (Applied Statistics). Vol. 51. n. 4. 375-392 (2002)

Anazawa, K., Yoshida, M.: Volatile contents in Japanese volcanic rocks: Application of multivariate analysis. Geochemical Journal. 28(4), 307-315 (1994)

Avery, M. R., Panter, K. S., Gorsevski, P. V.: Distinguishing styles of explosive eruptions at Erebus, Redoubt and Taupo volcanoes using multivariate analysis of ash morphometrics. Journal of Volcanology and Geothermal Research. Volume 332, 15 February 2017, Pages 1-13 (2017)

Cardellini, C., Chiodini, G., Frondini, F., Avino, R., Bagnato, E., Caliro, S., Lelli, M., & Rosiello A.: Monitoring diffuse volcanic degassing during volcanic unrests: the case of Campi Flegrei (Italy), Scientific Reports. 7, 6757(2017). doi:10.1038/s41598-017-06941-2

Chiodini, G., Avino R., Caliro S., & Minopoli C.: Temperature and pressure gas geoindicators at the Solfatara fumaroles (Campi Flegrei). Annals Geophysics. 54, 2 (2011). doi: 10.4401/ag-5002

Chiodini G., Caliro, S., Cardellini, C., Granieri, D., Avino, R., Baldini, A., Donnini M., Minopoli, C.: Long-term variations of the Campi Flegrei. Italy. volcanic system as revealed by the monitoring of hydrothermal activity. Journal of geophysical research. vol. 115. B03205 (2010). doi: 1029/2008JB006258

Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis. IEEE transactions on information theory. 36(5), 961-1005 (1990)

Deino, A. L., Orsi, G., De Vita, S., Piochi, M.: The age of the Neapolitan Yellow Tuff caldera forming eruption (Campi Flegrei caldera, Italy) assessed by 40Ar/39Ar dating method. Journal of Volcanology and Geothermal Research. 133, 157–170 (2004). https://doi.org/10.1016/S0377-0273(03)00396-2

Egozcue, J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barcelo-Vidal, C.: Isometric logratio transformations for compositional data analysis. Mathematical Geology. 35(3), 279-300 (2003)

Farge, M.: Wavelet transforms and their applications to turbulence. Annual review of fluid mechanics. 24(1), 395-458 (1992)

Filzmoser, P., Hron, K., Reimann, C.: Principal component analysis for compositional data with outliers. Environmetrics. 20(6), 621-632 (2009). doi: 10.1002/env.966

Gutenberg, B., Richter, C.F.: Seismicity of the Earth and Associated Phenomena, 2nd ed. Princeton, N.J.: Princeton University Press (1954)

Hernández-Antonio, A., Mahlknecht, J., Tamez-Meléndez, C., Ramos-Leal, J., Ramírez-Orozco, A., Parra, R., Ornelas-Soto, N., Eastoe, C. J.: Groundwater flow processes and mixing in active volcanic systems: the case of Guadalajara (Mexico*).* Hydrology and Earth System Sciences. 19, 3937–3950 (2015)

Isaia, R., D'Antonio, M., Dell'Erba, F., Di Vito, M., Orsi, G.: The Astroni volcano: the only example of closely spaced eruptions in the same vent area during the recent history of the Campi Flegrei caldera (Italy), Journal of Volcanology and Geothermal Research. 133(1), 171-192 (2004). https://doi.org/10.1016/S0377-0273(03)00397-4

Isaia, R., Marianelli, P., Sbrana, A.: Caldera unrest prior to intense volcanism in Campi Flegrei (Italy) at 4.0 ka B.P.: Implications for caldera dynamics and future eruptive scenarios, Geophysical Research Letters. 36, L21303 (2009). doi:10.1029/2009GL04513,1-6.

Jamieson. R. A., Baldini, J.U.L., Frappier, A.B., Müller, W.: Volcanic ash fall events identified using principal component analysis of a high-resolution speleothem trace element dataset. Earth and Planetary Science Letters. 426, 36–45 (2015)

Join, J.L., Coudray, J., Longworth, K.: Using Principal Components Analysis and Na/Cl Ratios to Trace Groundwater Circulation in a Volcanic Island: the Example of Reunion. Journal of Hydrology. 190 (1-2), 1-18 (1997)

Kumamoto, T., Tsukada, M., Fujita, M.: Multivariate Statistical Analysis for Seismotectonic Provinces Using Earthquake, Active Fault, and Crustal structure Datasets. Earthquakes, Tsunamis and Nuclear Risks. pp 31-41 (2016)

Steinhorst, K.J., Hodge, V.F., Guo, C., Farnham, M., Johannesson, K.H.: Geochemical and Statistical Evidence of Deep Carbonate Groundwater within Averlying Volcanic Rock Aquifers/aquitards of Southern Nevada, USA. Journal of Hydrology. 243, 254-271 (2001)

Torrence, C., Webster, P. J.: Interdecadal Changes in the ENSO-Monsoon System. American Meteorological Society. Vol. 12, 8 (1999)

Velasco-Tapia, F.: Multivariate Analysis, Mass Balance Techniques, and Statistical Tests as Tools in Igneous Petrology: Application to the Sierra de las Cruces Volcanic Range (Mexican Volcanic Belt). Hindawi Publishing Corporation, The Scientific World Journal, Volume 14 (2014)

Wiemer, S., Wyss, M.: Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. Bulletin of the Seismological Society of America. 90(4),859-869 (2000). doi: 10.1785/0119990114