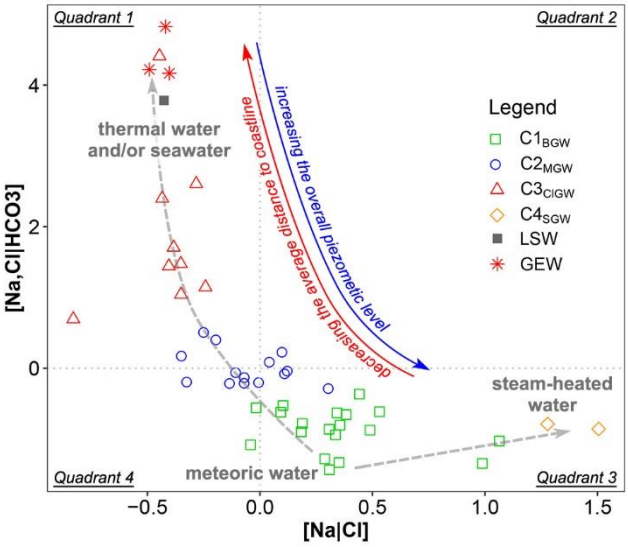
Highlights

- Ratios of HCO_3^-, SO_4^{2-} and Cl^- are proposed for performing hierarchical clustering
- Hierarchical clustering characterizes the main hydrogeochemical processes
- Bicarbonate–rich, chlorine–rich, sulfate–rich and mixed groundwaters are identified
- Groundwater geochemistry is investigated with compositional data analysis
- Compositional data analysis is useful for hydrogeochmical studies in volcanic aquifers

Graphical abstract



The main hydrogeochemical processes in Campi Flegrei volcanic aquifer, south Italy: Proposing anion ratios for hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry

Pooria Ebrahimi [a]*, Annalise Guarino [a], Vincenzo Allocca [a], Stefano Caliro [b], Rosario Avino [b], Emanuela Bagnato [b], Francesco Capecchiacci [b,c], Antonio Carandente [b], Carmine Minopoli [b], Alessandro Santi [b], Stefano Albanese [a]

[a] Department of Earth, Environmental and Resources Sciences, University of Naples Federico II, Naples 80126, Italy
[b] Istituto Nazionale di Geofisica e Vulcanologia sez. "Osservatorio Vesuviano", Naples 80124, Italy
[c] Department of Earth Sciences, University of Florence, Florence 50121, Italy

* Corresponding author: pooria.ebrahimi@unina.it, pooria.ebrahimi@gmail.com

Abstract

Comprehensive hydrogeochemical studies have been conducted in the Campi Flegrei volcanic aquifer since late 20th century due to the volcanic unrest. In the last decade, groundwaters were grouped based on the dominant anion species (i.e. bicarbonate, sulfate and chloride) to explain the general hydrogeochemical processes. In this article, 44 groundwater samples are collected from Campi Flegrei aquifer to geochemically and spatially capture the main characteristics of the groundwater body. The hierarchical clustering algorithm is then performed on proportion of bicarbonate, sulfate and chloride, and the optimum number of clusters are determined regarding the results of deep hydrogeochemical investigations published in the past. The collected samples are categorized in the following groups: (1) bicarbonate–rich groundwaters; (2) chlorine–rich groundwaters; (3) sulfate–rich groundwaters; and (4) mixed groundwaters. The first group is mainly derived from poor arsenic and lithium meteoric water, but there is significant thermal/seawater contribution in the second one. The interaction of bicarbonate–rich groundwaters and hydrothermal vapors (with high arsenic and hydrogen sulfide contents) give rise to the sulfate–rich groundwaters around Solfatara volcano. The mixed groundwaters are observed where the three main groundwater groups undergo a mixing process, depending on the hydrogeology of the volcanic aquifer. Contrary to bicarbonate– and sulfate–rich groundwaters, chlorine–rich and mixed groundwaters generally occur at low piezometric levels (approximately <1 m above sea level) near the coastline. The hierarchical cluster analysis provide more information about the volcanic aquifer, particularly when compositional data analysis is applied to study hydrogeochemistry of the homogeneous groundwater groups and to uncover the relationships between variables. Addressing compositional nature of data is recommended in the future studies for developing new tools that help deeper understanding of groundwater evolution in volcanic aquifers.

Keywords: Hierarchical cluster analysis; Groundwater evolution; Hydrothermal system; Solfatara; Compositional nature of data; Mixing process

1 Introduction

Geothermal activities occur in the areas of active volcanism, continental collision zones, continental rift systems associated with active volcanism, as well as continental rifts not associated with volcanoes (Chandrasekharam and Bundschuh, 2002; Moeck, 2014). A wide range of secondary processes affect chemical composition of the rising geothermal fluids from the reservoir to the surface. In coastal areas, thermal fluids are generally brackish to saline Na–Cl type because seawater alters the original chemical characteristics of the fluids (Dotsika, 2015). Different extents of mixing process between meteoric water, geothermal fluids and steam results in various hydrogeochemical signatures. Delineating these signals provides some information about evolution of the hydrothermal system, geothermal exploration and volcanic activity.

Campi Flegrei is one of the most active volcanic areas in the world, well-known for the striking ground movements, in which thermal energy is mainly released through diffuse degassing (Chiodini et al., 2001). Groundwater hydrogeochemistry was extensively investigated in the Campi Flegrei volcanic aquifer before 2007 (Aiuppa et al., 2006; Avino et al., 1999; Capaldi et al., 1992; Celico et al., 1992; Valentino et al., 1999; Valentino and Stanzione, 2003; 2004), but recent studies focused on specific sectors of the aquifer such as Bagnoli–Fuorigrotta Plain (Arienzo et al., 2015; De Vivo and Lima, 2008), Agnano Plain, Solfatara volcano (Bagnato et al., 2009) and Cuma (Allocca et al., 2018; Stellato et al., 2020). Noticeable groundwater heterogeneity likely reflects inhomogeneous water discharge, various residence times of groundwater in the aquifer and/or spatially variable contributions from deep thermal fluids (Aiuppa et al., 2006).

Since the beginning of this century, Campi Flegrei showed signs of volcanic unrest (Chiodini et al., 2010; 2012): changes in behavior of the volcanic system, composition of fumarolic gases and frequency of earthquake swarms. These observations are indicators of higher magmatic component and ground uplift which raised concerns of the scientific community about volcanic reawakening. The Campi Flegrei caldera and the surrounding environment have been highly populated since the ancient Greek times. In this region, cities (e.g., Naples, Pozzuoli and Cuma) form a seamless urban area with more than 2 million inhabitants (Giudicepietro et al., 2019). Hence, there is a need to better understand the present volcanic unrest and to be prepared for a possible emergency. To evaluate volcanic hazards, geochemical tracers provide unique information about spatial distribution and temporal rate of magma degassing along with activity of the magma chamber. However, hydrogeochemical assessment of volcanic aquifers situated in densely populated coastal areas is challenging because many factors affect groundwater chemistry.

Geochemical data are regularly reported as proportions such as weight percentages (wt%), milligrams per liter (mg/l) and micrograms per liter (µg/l). It means that such data subject to a constant sum because each component explain only a part of the whole composition. Karl Pearson introduced the problem of spurious correlation in this type of data (Pearson, 1987) and applying statistical analysis to compositional data has been an issue for more than a century. Compositional data analysis refers to a family of log-ratio transformations which was introduced by Aitchison (1982) and Egozcue et al. (2003) for converting the compositional data from their original sample space to an unrestricted real space. It has attracted attention of many researchers (Buccianti and Grunsky, 2014; Dominech et al., 2021; Ebrahimi et al., 2021) because the ratios between compositional parts are much more informative for understanding the complex geochemical data structure. To the best of our knowledge, it has not been used for studying hydrothermal systems and groundwater geochemistry in volcanic areas. The main objective of this work is taking advantage of the long history of hydrogeochemical studies in the study area to apply hierarchical cluster analysis (HCA) and compositional data analysis for investigating groundwater geochemistry, which might provide new tools for monitoring volcanic activity more accurately and mitigating risk of volcanic eruption more effectively.

2 Study area

The volcanic system of Campi Flegrei is located in an area of extensive tectonic activity on the Tyrrhenian coast, south Italy (Fig. 1). It covers an area of ca. 200 km^2 (Sellerino et al., 2019) with the maximum elevation of about 250 m above sea level. Distribution of the volcanic centers was controlled by NE-SW and NW-SE faults in the study area. The volcanic activity began before 80 ka (Neapolitan volcanoes; Scarpati et al., 2013) and the last eruption (Monte Nuovo) occurred in 1538 A.D. The majority of these volcanic eruptions were explosive (Vitale and Isaia, 2014). Campanian Ignimbrite (CI: 35 ka B.P.) and Neapolitan Yellow Tuff (NYT: 12 ka B.P.) eruptions led to formation of the present 12-km wide caldera (Aiuppa et al., 2006). Varying from trachybasalts to phonolitic and peralkaline trachytes in composition (Armienti et al., 1983; Di Girolamo, 1978), the volcanic products of Campi Flegrei can be considered as a part of the potassium–rich Roman comagmatic province in central-southern Italy (Peccerillo, 1985; Washington, 1906). Recent active continental sediments and volcanic deposits younger than 12 ka covered the area and the older outcrops occurred around the NYT and CI caldera rims (Fig. 1). The fumaroles in the Pozzuoli Bay and Solfatara volcano along with the hot springs and steam-heated pools indicate intense hydrothermal activity in the study area (Aiuppa et al., 2006).

Since the middle of the last century, the caldera experienced a long-term bradyseismic crisis associated with temporal injections of CO_2–rich magmatic fluids at the bottom of the hydrothermal system (Caliro et al., 2007). This resulted in periodic ground uplifts and seismic activities with a maximum total uplift of approximately 4 meters from 1983 to 1984 (Del Gaudio et al., 2010). It followed an episode of subsidence until 2004–2005 when another uplift phase began. From 2012 to 2013, deformation rate was accelerated (D'Auria et al., 2015; Trasatti et al., 2015) due to magma emplacement at shallow depth (D'Auria et al., 2015). The recent uplift phase is still ongoing leading to significant variation in degassing rate, fumarolic composition and shallow earthquakes (Cardellini et al., 2017; Chiodini et al., 2010; Giudicepietro et al., 2019). Variations in pH, ammonia and/or SO_4^{2-}/Cl^- ratio in pools and wells around the Solfatara crater (Fig. 1) from 1990 to 1999 revealed that this area is the best site for monitoring the changes in the deep hydrothermal system and the activities of the underlying magma chamber (Valentino and Stanzione, 2004).

Various origins were proposed for the thermal waters in Campi Flegrei: (1) Seawater or brines probably exist at depth that will be heated by the rising magmatic fluids (Dall'Aglio et al., 1972); (2) Local meteoric and deep hot marine components are mixed with different proportions (Baldi et al., 1975; Cortecci et al., 1978); (3) Thermal water and fumarolic steam mostly represent the meteoric origin (Bolognesi et al., 1986); (4) Thermal water and fumarolic steam are composed of a mixture of hot deep steam and local shallow water (Ghiara et al., 1988); (5) A shallow reservoir (depth <2000 m) with seawater and local meteoric water origin, and a deep reservoir (depth >2000 m) containing magmatic-meteoric hypersaline fluids feed the geothermal system of Campi Flegrei (Caprarelli et al., 1997). High As concentrations in Na–Cl brines originates from prolonged water–rock interactions under reservoir temperature, fO_2 and fH_2S conditions, and the buffering effect of arsenopyrite–pyrite–pyrrhotite mineral assemblage. The brines are then diluted by As-depleted meteoric-derived groundwaters during their ascent towards the surface and give rise to mixed waters with intermediate to low arsenic contents (Aiuppa et al., 2006). Regarding the gaseous composition of Solfatara fumaroles, Caliro et al. (2007) proposed two thermobarometric signatures at: (1) around 360°C indicating the deep zone of hydrothermal system where CO_2–rich magmatic gases flash the hydrothermal fluid of meteoric origin (with fixed CO_2 fugacity due to fluid–rock interactions at high temperature) and form a gas plume; and (2) about 200–240°C representing re-equilibration of the reactive species (i.e., H_2 and CO) in the rising gas plume. Extensive interaction of rocks with As-bearing

hydrothermal steam results in intermediate to high As in steam-heated groundwater (Aiuppa et al., 2006).

Valentino and Stanzione (2003) collected several water samples from 13 stations during 1994, considered the average major ion contents (literature data were included in some cases) and identified four end-members in Campi Flegrei: (1) alkali–chloride, highly saline waters (Stufe di Nerone well and Stufe di Nerone spring); (2) alkali-bicarbonate, less saline waters (Agnano spring and the waters from Quarto plain); (3) alkali–bicarbonate–sulfate waters (a sampling station in Quarto plain); and (4) alkaline earth–sulfate, acidic waters (Pisciarelli Grande and Pisciarelli Piccola). They proposed that the groundwaters from Pozzanghera Agnano Terme, Hotel Tennis, Terme Puteolane and Damiani are mixture of the end-members. However, Aiuppa et al. (2006) used the percent meq/l of HCO_3^-, SO_4^{2-} and Cl^-, defined the prevalent anion species and grouped 64 samples in bicarbonate–, chlorine– and sulfate–rich groundwaters. The latter classification can be considered as the generalized version of the former, being beneficial when investigating a large number of samples.

Hydrogeologically, Campi Flegrei volcano is located at the southwestern boundary of Campania Plain. Little is known about the hydrologic system to date. A very complex stratigraphy was indicated by the wells drilled up to ca. 3 km depth for geothermal exploration in Mofete, S. Vito and Agnano areas (AGIP, 1987; Rosi and Sbrana, 1987). Tuff, lava and subordinate sedimentary rocks are followed by their thermometamorphic equivalents at ~2 km depth. The layers are impervious in more than a few hundred meters depth (also due to the self-sealing processes originating from circulation of hydrothermal fluids), except for the fractured zones (Capaldi et al., 1992). The complex subsurface geology significantly influences the shallow hydrologic system and leads to presence of superimposed aquifers in the study area. These aquifers, at a regional scale, can be considered as a unique aquifer due to horizontal discontinuities and geometrical variability of the lithotypes (Celico et al., 1987). The high piezometric level near Quarto plain (approximately 20 m a.s.l.; Fig. 1) is feasibly associated with deep fluids upwelling (Piochi et al., 2014) resulting in a pseudo-radial groundwater flow pattern at a regional scale. Accordingly, the groundwater flows towards the coastline in the southern, southeastern and western sectors, and towards the Campanian Plain aquifer in the northern sector (Fig. 1) (Celico et al., 1992; De Vita et al., 2018). Before reaching the sea, groundwater naturally emerges at the Averno volcanic lake, Lucrino, Fusaro and Miseno coastal lakes (Fig. 1) together with some thermal springs in Agnano plain and along the coast (Sellerino et al., 2019). The annual water-table fluctuation varies from 0.12 to 0.95 m, being the lowest in Agnano Plain. The time lag between precipitation and increase of water-table level ranges from 1 to 2 months in Quarto plain although there is an immediate correspondence in Agnano Plain. High water-table level was observed in the thermal wells of Damiani and Hotel Tennis (Fig. 1) 5 to 6 months after the timespan of major rainfall (Valentino and Stanzione, 2004).

3 Materials and methods

3.1 Sampling and chemical analysis

Majority of the published hydrogeochemical data in Campi Flegrei refer to specific monitoring points which are not suitable for the purpose of the present study because the samples do not cover the whole study area. In May–June 2002, sixty-four groundwater samples were collected and the analytical results were published (Aiuppa et al., 2006), but location and descriptive ID of each sampling point are not available. Therefore, in this study, a total of 44 water samples are collected in May 2019 from wells and springs mostly located in the NYT caldera (Fig. 1, Table S1). Alkalinity, electric conductivity, pH and temperature are measured with portable instruments in the field. The water

samples are filtered with 0.45 μm filters and collected in 30 ml high-density polyethylene (HDPE) bottles for chemical analyses. One aliquot is acidified (1%) with Suprapure 36% HCl for determination of the major cations, whereas the aliquot for minor and trace elements is acidified (1%) with Suprapure 63% HNO_3. Samples are analyzed for Li, Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, SO_4^{2-}, Cl^{-}, NO_3^{-} and F^{-} by ion chromatography standard technique, using a Dionex ICS3000 system at the laboratory of fluid geochemistry at the INGV-Osservatorio Vesuviano, Napoli. Measurement accuracy is better than ±5%, obtained by analyzing certified standard solutions, and the detection limits are better than 0.1 mg/l. Arsenic and boron contents are detected using inductively coupled plasma–optical emission spectrometry (ICP–OES) at the Department of Earth, Environmental and Resources Sciences in the University of Naples Federico II, following 0.45-μm filtration and HNO_3 acidification of another aliquot in the field. Measurement accuracy is better than ±5%, obtained by analyzing certified standard solutions, and the detection limits are 3.0 μg/l.

3.2 Statistical analysis

3.2.1 Hierarchical cluster analysis (HCA)

Most natural datasets are composed of several unknown subpopulations with specific characteristics (e.g., average, median and standard deviation). HCA is an exploratory statistical approach for uncovering these subgroups and creating homogeneous clusters in a dataset. The samples occurring in a cluster tend to have the highest possible similarity to each other, but the differences between the clusters are as large as possible. However, like factor analysis and principal component analysis, it cannot itself be a "statistical proof" of a certain relationship between samples and the corresponding groups (Reimann et al., 2008). In HCA, a distance matrix is computed which explains the similarity degree between observations and each observation is initially considered as an isolated group. Then, the two most similar ones are joined using an agglomeration technique and this process continues until all observations are pooled in a single group. The distances between clusters are recalculated at each stage depending on the selected method (Filzmoser et al., 2018; Keshavarzi et al., 2015; Van den Boogaart and Tolosana-Delgado, 2013). Euclidean measure is chosen to obtain the distance between the individuals and the McQuitty method (McQuitty, 1966) is applied as the agglomeration technique in the present article. The R package "stats" (R Core Team, 2021) is utilized to perform hierarchical clustering.

Variables with very different levels (e.g., major vs. trace elements) should not undergo HCA without proper data transformation and standardization, otherwise the variable with the highest variance affects the outcome noticeably. Including or excluding only one variable can also lead to completely different hierarchical clustering results (Reimann et al., 2008). In this study, ratios of bicarbonate, sulfate and chloride are considered for cluster analysis (Fig. 2) because the major ions could show the main processes regulating groundwater composition in Campi Flegrei (Aiuppa et al., 2006):

$$Cl^{-}\ \text{ratio} = \frac{Cl^{-}}{HCO_3^{-}+SO_4^{2-}+Cl^{-}} \quad (1)$$

$$SO_4^{2-}\ \text{ratio} = \frac{SO_4^{2-}}{HCO_3^{-}+SO_4^{2-}+Cl^{-}} \quad (2)$$

$$HCO_3^{-}\ \text{ratio} = \frac{HCO_3^{-}}{HCO_3^{-}+SO_4^{2-}+Cl^{-}} \quad (3)$$

where all values are in mg/l. Presence of outliers and variance of the above-mentioned ratios are double-checked prior to hierarchical clustering (Fig. 2). Data outliers likely affect the distance

measures and distort the data structure, but they need to be appropriately accommodated because they may signal some unexpected behavior in the hydrothermal system (Reimann et al., 2008).

3.2.2 Compositional data analysis

Majority of geochemical data are concentrations of constituents, relative in nature, depend on one another and sum up to a constant. These data are called compositional data and they are restricted to the positive part of the real sample space. Hence, their algebraic-geometric structure (i.e. the Aitchison geometry) is different from that of Euclidian geometry in real space. Applying traditional interpretive methods such as scatterplots and correlation analysis to data on the simplex can be misleading (Pearson, 1987; Kynčlová et al., 2017; Reimann et al., 2017). Compositional data analysis refers to the statistical approach that respects geometry of compositional data and helps for interpreting the relationships between relative components. The additive log-ratio (alr), centred log-ratio (clr) and isometric log-ratio (ilr) transformations were introduced by Aitchison (1982) and Egozcue et al. (2003) for solving the problem by converting a composition to a real vector prior to data elaboration in unconstrained conditions.

3.2.2.1 The isometric log-ratio (ilr)-ion plot

The Piper diagram was introduced in 1944 to classify different water types, characterize mixing of water masses and determine the chemical reactions regulating water chemistry in a set of water samples under investigation (Piper, 1944). Although application of the diagram has obviously been beneficial in the last eight decades, it has been challenging in some circumstances. For instance, distinguishing between waters mainly composed of Ca^{2+}, Mg^{2+}, SO_4^{2-} and Cl^- is difficult particularly for large datasets because of the structure of the diamond-shaped field in Piper diagram. Another issue is clustering the points along the boundaries and in the corners of the plot when water is concentrated in a particular pair of cation and anion. In order to overcome the above-mentioned shortcomings and improve the robustness of the Piper diagram, Shelton et al. (2018) used compositional data analysis and calculated four isometric log-ratios for the major chemical species of the original diagram as follows:

$$z_1 = \sqrt{\frac{2}{3}} \ln \frac{\sqrt{[Ca^{2+}][Mg^{2+}]}}{[Na^{+}+K^{+}]} \quad (4)$$

$$z_2 = \frac{1}{\sqrt{2}} \ln \frac{[Ca^{2+}]}{[Mg^{2+}]} \quad (5)$$

$$z_3 = \sqrt{\frac{2}{3}} \ln \frac{\sqrt{[Cl^{-}][SO4_4^{2-}]}}{[HCO_3^{-}+CO_3^{2-}]} \quad (6)$$

and

$$z_4 = \frac{1}{\sqrt{2}} \ln \frac{[Cl^{-}]}{[SO_4^{2-}]} \quad (7)$$

where all concentrations are in meq/l. The ilr-ion plot is then generated as a four-panel scatterplot (Fig. 2). Shelton et al. (2018) compared the Piper diagram and the proposed alternative diagram using multiple synthetic and real datasets and indicated that the latter offers a more in-depth assessment of water chemistry. The ilr-ion plot could be produced through the R code provided by Shelton et al. (2018).

3.2.2.2 Scatterplots

Generating scatterplots by raw or log-transformed data probably results in misinterpretation because of ignoring the compositional nature of the data (Reimann et al., 2017). Symmetric coordinates, a special case of isometric log-ratio transformation, is an option to solve this data problem by capturing all the relative information regarding the two compositional parts of interest (Kynčlová et al., 2017). In this study, symmetric coordinates and the frequently used classical log-transformation are chosen to investigate the importance of data closure (i.e. the constant sum constraint) (Fig. 2). Previous studies indicated that the former characterizes the relationships between elements better (Somma et al., 2021). The R package "robCompositions" (Filzmoser et al., 2018) is used for computing the symmetric coordinates in Figs. 8 and 9, but the readers interested in symmetric coordinates are referred to Kynčlová et al. (2017) for further information. Linear regression model is finally applied via the R package "ggpmisc" (Aphalo, 2021) and the coefficient of determination (R^2) is obtained for the selected parameters.

In addition, two isometric log-ratios were constructed with Na^+, Cl^- and HCO_3^- concentrations following the sequential binary partitioning procedure (Egozcue and Pawlowsky-Glahn, 2005). These coordinates are also called balances. Sodium and chloride are the major components of Na–Cl brines and seawater whilst bicarbonate is characteristic of meteoric water and CO_2–rich hydrothermal waters. Generating the following balance can describe the sources of Na^+ and Cl^- in comparison with those of HCO_3^-:

$$[Na^+, Cl^- | HCO_3^-] = \sqrt{\frac{2}{3}} \ln \frac{\sqrt{[Na^+][Cl^-]}}{[HCO_3^-]} \quad (8)$$

where all values are in mg/l. The second balance also helps to assess relative variation of Na^+ and Cl^- and detect the samples similar to Na–Cl brines and seawater:

$$[Na^+ | Cl^-] = \frac{1}{\sqrt{2}} \ln \frac{[Na^+]}{[Cl^-]} \quad (9)$$

where all concentrations are in mg/l. The R code for generating a scatterplot based on Eqs. 8 and 9 is available in supplementary material.

4 Results and discussion

In this section, groundwater samples are categorized in three groups based on the dominant anion species and in four groups regarding the results of hierarchical cluster analysis (Table 1). To avoid confusion, the abbreviations (i.e. BGW, ClGW and SGW which stand for bicarbonate–, chlorine– and sulfate–rich groundwaters, respectively) are exclusively used for discussing the dominant anion groups. However, the HCA groups are *mainly* referred to bicarbonate–rich, mixed, chlorine–rich and sulfate–rich groundwaters ($C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$, respectively).

4.1 An overview of groundwater composition

Na–Cl and Na–HCO_3 are the main groundwater types in Campi Flegrei (Table 1). However, Na–SO_4 is exclusive to the groundwaters around Solfatara crater with the highest temperatures (>79°C). On average, Na–Cl groundwaters are 10°C warmer than Na–HCO_3 groundwaters. The average contents of cations and anions follow the orders of Na^+ (1567.8 mg/l) > K^+ (172.8 mg/l) > Ca^{2+} (143.9 mg/l) > Mg^{2+} (51.9 mg/l) and Cl^- (2278.7 mg/l) > HCO_3^- (757.6 mg/l) > SO_4^{2-} (409.7 mg/l) > NO_3^- (56.3 mg/l) > F^- (5.0 mg/l), respectively. Lithium, arsenic and boron show the highest robust coefficient of variation (CVR) which probably suggest their unique source. The pH values range from 5.7 in Na–Cl groundwater types to 8.6 in Na–HCO_3 groundwaters revealing the acidic to alkaline nature

of the groundwaters under investigation. Nonetheless, electric conductivity indicates a reverse trend, decreasing from Na–Cl groundwaters to Na–HCO_3 groundwaters (Table 1).

Regarding the major anion groups (Table 1), BGW (T = 16.7 to 53.4 ^{o}C; EC = 350 to 6340 μS/cm) mostly represent the meteoric-derived waters which can be supported by the isotopic investigations of Caprarelli et al. (1997). Mixing these groundwaters and Na–Cl brines (i.e. the chemically and isotopically modified seawater due to heating, boiling and interacting with rocks; Caprarelli et al. (1997)) leads to ClGW (T = 16.4 to 77 ^{o}C; EC = 795 to 37000 μS/cm). On the other hand, the interaction between BGW and the vapor phase form SGW with high temperature (>79^{o}C) and intermediate electric conductivity (3400 to 4480 μS/cm). Chiodini et al. (2001, 2003) reported high $H_2S_{(g)}$ in the condensing hydrothermal vapors. Oxidizing the H_2S by atmospheric gases results in elevated sulfate concentration in the groundwater around Solfatara crater.

4.2 Detection of homogeneous water groups

Many data outliers are detected in Cl^- and SO_4^{2-} boxplots and variance of chloride is two orders of magnitude higher than those of bicarbonate and sulfate (Figs. S1a, b and c). However, computing anion ratios alleviate these data problems for applying clustering algorithm (Figs. S1d, e and f). The collected samples can be categorized in bicarbonate–rich, sulfate–rich, chlorine–rich and mixed groundwaters based on Fig. 3. On average, the first two groups are spatially located far from the coastline, where the piezometric level is >1 m a.s.l. (Figs. 1 and 3 and Table S2). As the groundwater body flows towards the coast, mixed and chlorine-rich groundwaters become dominant. It is worth mentioning that the water from Agnano spring (W40 in Tables 1 and S1) is likely affected by magmatic CO_2 and buffered by water–rock interactions (Valentino and Stanzione, 2004), but it is grouped with the meteoric waters ($C1_{BGW}$) due to the similar bicarbonate, sulfate and chloride contents. Nevertheless, compared with the dominant ion groups, hierarchical cluster analysis divides ClGW into (Fig. 3):

(1) *the mixed groundwaters ($C2_{MGW}$)*. Valentino and Stanzione (2003) introduced the water samples collected in Pozzanghera Agnano Terme, Hotel Tennis, Terme Puteolane and Damiani as a mixture of the end-members. These samples occur in this group and consequently no predominant source (or a mixed source) is expected for the whole group. It is worth mentioning that bicarbonate is the dominant anion in Puteolane (W39 in Fig. 3), but HCA results express that it underwent a mixing process which is consistent with the measurements in the late 20th century (Table S3).

(2) *the chlorine–rich groundwaters ($C3_{ClGW}$).* The groundwaters around Stufe di Nerone well and Stufe di Nerone spring, and the groundwaters in Bagnoli coastal plain are present in this group. In Stufe di Nerone well, the deep geothermal component is hot 'mature' seawater modified through water-rock interactions. Ammonium contents in waters from Stufe di Nerone well and Stufe di Nerone spring also suggest a mixing process between deep geothermal and shallow meteoric components (Valentino and Stanzione, 2004). In addition, rise of hydrothermal fluids along fractures or faults was reported in the Bagnoli plain by Celico et al. (2001) and De Vivo and Lima (2008). Hence, the groundwaters with predominant thermal and/or seawater component likely present in this group.

Although the mixed and bicarbonate–rich groundwaters are the highest similar groups, because they join together in Fig. 3, the statistically significant difference in average Cl^- and HCO_3^- ratios between them reveals their effective clustering (Fig. 4). It is noteworthy that HCA considers the multivariate

data space by utilizing the distances between observations. Significance of difference in the average ratios between the other combinations of clusters is not assessed due to their few observations and limitation of the bootstrapping approach. Sulfate–rich groundwaters are the next group that merges these two at a higher level of the dendrogram (Fig. 3), probably showing a relatively stronger fingerprint of meteoric water in composition of the groundwaters in the study area, except for the chlorine–rich groundwaters.

Chemical signature of the four groups of HCA versus the three groups of dominant anion species are compared in Fig. 5. It is evident that the percentages of bicarbonate and sulfate (left panels in Figs. 5a and c) decrease from the meteoric end-member ($C1_{BGW}$) to the thermal/seawater end-member ($C3_{ClGW}$) although chloride percentage, arsenic, boron, lithium, electric conductivity and temperature (left panels in Figs. 5b, d, e, f, g and h) show a reverse trend. Values of the selected parameters in the mixed groundwaters range between the corresponding values in bicarbonate– and chlorine–rich groundwarers which confirm their hybrid nature. This information cannot be obtained from categorizing the samples based on the dominant anion species (i.e. the right panels in Fig. 5). The sulfate–rich groundwaters have the lowest Cl^- (%), but the highest SO_4^{2-} (%) and temperature (^oC) (Figs. 5b, c and h).

4.3 Compositional data analysis and hydrogeochemistry of volcanic aquifers

4.3.1 The hidden information in the major ions data

Composition of the four groundwater groups (i.e. $C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$) is compared with chemistry of the geothermal fluid and local seawater through Piper diagram (Fig. 6) and ilr-ion plot (Fig. 7). A quick look at these two graphical approaches for analyzing water chemistry clarifies the advantages of utilizing the compositional alternative of the original Piper diagram. Regardless of the different isometric log-ratios in Fig. 7, chlorine–rich groundwaters are not condensed in the corners of the ilr-ion plot and their chemical variation can be assessed in detail.

Excluding sulfate–rich groundwaters, the positive correlation between the samples is better demonstrated in Fig. 7b. The bicarbonate–rich grounwaters show the lowest Cl^-/SO_4^{2-} together with Cl^- and SO_4^{2-} against $HCO_3^- + CO_3^{2-}$. Values of these ratios gradually increase in the mixed and chlorine–rich groundwaters confirming higher contribution of seawater and/or Na–Cl brines. It is noteworthy that the greatest variation of $[Cl^-, SO_4^{2-} | HCO_3^- + CO_3^{2-}]$ is observed in chlorine–rich groundwaters whilst significant changes of $[Cl^- | SO_4^{2-}]$ is evident in bicarbonate–rich groundwaters (Fig. 7b). It might show the complex situation in the study area which cannot be detected in the original Piper diagram (Fig. 6c).

Although it is difficult to distinguish all water groups in cation trilinear plot of the Piper diagram (Fig. 6b) due to significant data clustering, the data points are scattered in the ilr-ion plot (Fig. 7c) which allows to decipher the underlying hydrogeochemical interactions. Seawater indicates the smallest Ca^{2+}/Mg^{2+} value. The ratio increases with decreasing Ca^{2+} and Mg^{2+} versus $Na^+ + K^+$ in chlorine–rich grounwaters (Fig. 7c). Neoformation of Mg-bearing clay minerals and dissolution of Na–rich vitreous volcanic rocks were previously proposed as the geochemical processes regulating the evolutionary trend of seawater (Valentino and Stanzione, 2003).

The lower-right panel of the ilr-ion plot (Fig. 7d) mimics the diamond-shaped field of the Piper diagram (Fig. 6a). Even though concentrations of Cl^- and SO_4^{2-} are equal to $HCO_3^- + CO_3^{2-}$in mixed groundwaters, bicarbonate–rich groundwaters are enriched in bicarbonates. In accordance with

seawater and thermal fluids, chlorine–rich groundwaters contain higher chloride and sulfate contents in comparison with bicarbonates. Further, the ratio of Ca^{2+} and Mg^{2+} versus $Na^{+} + K^{+}$ represent the following decreasing order (Fig. 7d): $C1_{BGW}$ > $C2_{MGW}$ > $C3_{ClGW}$.

In the ilr-ion plot, Fig. 7a provides interesting information about the hydrothermal system. Most of bicarbonate–rich groundwaters are composed of equal chloride and sulfate contents. Their mixture with hydrothermal/seawater results in increasing chloride, but steam-heating leads to elevated sulfate content. Calcium concentration is higher than magnesium level in majority of the samples (Fig. 7a).

4.3.2 The relationship between selected variables

In order to present the importance of considering all relevant information in compositional data, the strength of relationships between Na^{+}– Cl^{-}, B–Cl^{-}, and B–Na^{+} are evaluated via symmetric coordinates and log-transformed values (Fig. 8). In the groundwaters with significant contribution of seawater or the hydrothermal fluids, a higher correlation coefficient (i.e. a greater R^2 value) between the selected pairs of variables is expected. This is confirmed in Fig. 8 because the weakest and the strongest relationships exist between the pairs of variables in bicarbonate– and chlorine–rich groundwaters, respectively. The mixed groundwaters represent a moderate R-squared value. It is noteworthy that the R^2 values are higher in Figs. 8b, d and f due to spurious correlation. For instance, approximately 80% ($p<0.05$) of sodium variation is explained by chloride in the groundwaters mainly derived from meteoric water ($C1_{BGW}$). It reduces to 45% ($p<0.05$) when symmetric coordinates of Na^{+} and Cl^{-} are utilized for linear regression model which seems more reliable regarding the origin of groundwater (Figs. 8a and b). Contrary to Fig. 8f, the B–Cl^{-} relationship in chlorine–rich groundwaters is not statistically significant in Fig. 8e and deserves further evaluation with a larger dataset.

Owing to the conservative geochemical behavior of Li, it can be a potential tracer of the ascending hot deep brines (Giggenbach, 1991). Furthermore, elevated As concentration in Cl-bearing high temperature springs probably explains prolonged water–rock–magmatic gas interactions in reservoir conditions (Ballantyne and Moore, 1988). The Li–Cl^{-} and As–Cl^{-} relationships were investigated by Aiuppa et al. (2006) to explore geochemical evolution of groundwater in the study area. Hence, the compositional counterparts (i.e. symmetric coordinates) were utilized to assess the relationships in the true data structure. Lithium and Arsenic evidently increase with respect to Cl^{-} from meteoric derived waters towards thermal waters (Fig. 9). The chlorine–rich groundwaters can consequently be considered as the deep reservoir brine partially diluted with meteoric water. Depending on hydrogeology of the volcanic aquifer, Na–Cl brines may be noticeably diluted with poor Li and As meteoric waters and give rise to mixed groundwaters. Seawater contribution is likely another process explaining the deviation from the mixing trend between the bicarbonate–rich and chlorine–rich groundwaters (Fig. 9). Arsenic enrichment around Solfatara crater (the sulfate–rich groundwaters spatially occur in this area) is indicative of significant As in the vapor phase, condensing at shallow depth and giving rise to steam-heated groundwaters. It is in agreement with almost high As values in the fumarolic condensates of Solfatara (As∼3000 µg/l in the alkaline condensate; Aiuppa et al., 2006).

Plotting the isometric log-ratios generated using Na^{+}, Cl^{-} and HCO_3^{-} evidently represents geochemical evolution of the groundwater (Fig. 10). The bicarbonate–rich and sulfate–rich groundwaters occur in quadrant 3 of the graph ($[Na^{+}|Cl^{-}] > 0$ and $[Na^{+}, Cl^{-}|HCO_3^{-}] < 0$), but the chlorine–rich groundwaters lie in quadrant 1 ($[Na^{+}|Cl^{-}] < 0$ and $[Na^{+}, Cl^{-}|HCO_3^{-}] > 0$). Nevertheless, the mixed groundwaters occur around intersection of the grey dotted lines indicating that Na^{+} is equal to Cl^{-} and the ratio of sodium and chloride against bicarbonate is one. The overall peizometric level (Fig. 1) and the average distance to coastline (Fig. 3; Table S2) decrease as moving

from quadrant 3 to quadrant 1 (Fig. 10). It verifies that the bicarbonate-rich meteoric water flows towards the sea, interacts with the chlorine–rich thermal/seawater end-member and gives rise to the mixed groundwaters. Chlorine–rich groundwaters are characterized by the highest variation of Na^+ and Cl^- versus HCO_3^- and the lowest variation of Na^+/Cl^-. Their Na^+/Cl^- values are similar to those of local seawater and geothermal water, increasing as contribution of meteoric-derived water and water-rock interactions increases. Low ratios of sodium and chloride against bicarbonate in steam-heated waters suggest their meteoric origin (Valentino and Stanzione, 2004).

Conclusion

Campi Felgrei is an active volcano that has been investigated for several decades to understand the history of volcanic activity and monitor volcanic unrest. Hydrogeochemical studies are one of the important lines of research that help for characterizing risk of eruption. The previous studies showed that hydrogeochemical processes in the volcanic aquifer could be preliminarily understood using the dominant anion species (i.e. bicarbonate, chloride and sulfate). In the present article, major ions, Li, As, B, pH, electric conductivity and temperature are determined in 44 groundwater samples that are spatially distributed over the whole study area. Ratios of the above-mentioned anions are considered to run HCA, the representative number of clusters is determined regarding the detailed hydrogeochemical investigations in the study area and four groups are identified: (1) bicarbonate–rich groundwaters; (2) chlorine–rich groundwaters; (3) sulfate–rich groundwaters; and (4) mixed groundwaters. The relationships between variables in each group is then explored using compositional data analysis instead of the classical methods. Compared with the Piper diagram, the ilr-ion plot represents variations in major ions more clearly. In addition, the R-squared values obtained from linear regression model do not show spurious correlation when respecting the compositional nature of geochemical data. The scatterplot generated via the isometric log-ratios constructed by Na^+, Cl^- and HCO_3^- also indicate that the bicarbonate-rich meteoric water flows towards the sea, interacts with the chlorine–rich thermal/seawater end-member and gives rise to the mixed groundwaters. The rising hydrothermal vapor, however, increases sulfate and arsenic contents of the groundwater body around Solfatara crater. This article revealed some advantages of compositional data analysis, but applying this technique is recommended in future to highlight its potentials for mitigating volcanic risk more effectively.

Acknowledgement

The authors would like to thank Salvatore Dominech from School of Ocean and Earth Science at Tongji University (Shanghai, China) and Antonio Aruta from Department of Earth, Environmental and Resources Sciences at the University of Naples Federico II (Naples, Italy) for sharing ideas on application of compositional data analysis. We wish to express our gratitude to the associate editor and the anonymous reviewers for their constructive comments.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- AGIP, 1987, Modello geotermico del sistemaflegreo. Sintesi, SEG-MESG Internal Report, 23 pp.
- Aitchison, J., 1982. The statistical analysis of compositional data. Journal of the Royal Statistical Society: Series B (Methodological), 44(2), 139-177.
- Aiuppa, A., Avino, R., Brusca, L., Caliro, S., Chiodini, G., D'Alessandro, W., Favara, R., Federico, C., Ginevra, W., Inguaggiato, S., Longo, M., Pecoraino, G., Valenza, M., 2006. Mineral control of arsenic content in thermal waters from volcano-hosted hydrothermal systems: insights from island of Ischia and Phlegrean Fields (Campanian Volcanic Province, Italy). Chemical Geology, 229(4), 313-330.
- Allocca, V., Coda, S., De Vita, P., Di Rienzo, B., Ferrara, L., Giarra, A., Mangoni, O., Stellato, L., Trifuoggi, M., Arienzo, M., 2018. Hydrogeological and hydrogeochemical study of a volcanic-sedimentary coastal aquifer in the archaeological site of Cumae (Phlegraean Fields, southern Italy). Journal of Geochemical Exploration, 185, 105-115.
- Aphalo, P. J., 2021. ggpmisc: Miscellaneous Extensions to 'ggplot2'. R package version 0.3.9. https://CRAN.R-project.org/package=ggpmisc
- Armienti, P., Barberi, F., Bizojard, H., Clocchiatti, R., Innocenti, F., Metrich, N., Rosi, M., Sbrana, A., 1983. The Phlegraean Fields: magma evolution within a shallow chamber. Journal of Volcanology and Geothermal Research, 17(1-4), 289-311.
- Arienzo, M., Allocca, V., Manna, F., Trifuoggi, M., Ferrara, L., 2015. Effectiveness of a physical barrier for contaminant control in an unconfined coastal plain aquifer: the case study of the former industrial site of Bagnoli (Naples, southern Italy). Environmental monitoring and assessment, 187(12), 761.
- Avino, R., Capaldi, G., Pece, R., 1999. Radon in active volcanic areas of Southern Italy. Il nuovo cimento C, 22(3-4), 379-386.
- Bagnato, E., Parello, F., Valenza, M., Caliro, S., 2009. Mercury content and speciation in the Phlegrean Fields volcanic complex: evidence from hydrothermal system and fumaroles. Journal of Volcanology and Geothermal Research, 187(3-4), 250-260.
- Baldi, P., Ferrara, G.C., Panichi, C., 1975. Geothermal research in western Campania, southern Italy: chemical and isotopic studies of thermal fluids in the Campi Flegrei. In: Proc. 2nd U.N. Symp. on Development and Use of Geothermal Resources, San Francisco, Vol. 1, pp. 687–697.
- Ballantyne, J. M., Moore, J. N., 1988. Arsenic geochemistry in geothermal systems. Geochimica et Cosmochimica Acta, 52(2), 475-483.
- Bolognesi, L., Noto, P., Nuti, S., 1986. Studio chimico ed isotopico della Solfatara di Pozzuoli: ipotesi sull'origine e sulle temperature profonde dei fluidi. Rendiconti della Societa Italiana di Mineralogia e Petrologia, 41, 281–295.
- Buccianti, A., Grunsky, E., 2014. Compositional data analysis in geochemistry: Are we sure to see what really occurs during natural processes?. Journal of Geochemical Exploration, 1-5. https://doi.org/10.1016/j.gexplo.2014.03.022
- Caliro, S., Chiodini, G., Moretti, R., Avino, R., Granieri, D., Russo, M., Fiebig, J., 2007. The origin of the fumaroles of La Solfatara (Campi Flegrei, south Italy). Geochimica et Cosmochimica Acta, 71(12), 3040-3055.
- Capaldi, G., Pece, R., Veltri, C., 1992. Radon variation in groundwaters in the Campi Flegrei Caldera (southern Italy) during and after the 1982–1984 bradyseismic crisis. Pure and Applied Geophysics, 138(1), 77-93.
- Caprarelli, G., Tsutsumi, M., Turi, B., 1997. Chemical and isotopic signatures of the basement rocks from the Campi Flegrei geothermal field (Naples, southern Italy): Inferences about the origin and evolution of its hydrothermal fluids. Journal of Volcanology and Geothermal Research, 76(1-2), 63-82.
- Cardellini, C., Chiodini, G., Frondini, F., Avino, R., Bagnato, E., Caliro, S., Lelli, M., Rosiello, A., 2017. Monitoring diffuse volcanic degassing during volcanic unrests: the case of Campi Flegrei (Italy). Scientific Reports, 7(1), 1-15.
- Celico, P., De Gennaro, M., Pagano, D., Ronca, A., Stanzione, D., Vallario, A., 1987. Idrogeologia e idrogeochimica dei Campi Flegrei: relazioni tra chimismo delle acque e idrodinamica sotterranea, Atti Convegno "Bradisismo e fenomeni connessi", Napoli, 19-20 gennaio 1987, 94-100.

- Celico, P., Dall'Aglio, M., Ghiara, M. R., Stanzione, D., Brondi, M., 1992. Geochemical monitoring of the thermal fluids in the phlegraean fields from 1970 to 1990. Bollettino della Società geologica italiana, 111(3-4), 409-422.
- Celico, F., Esposito, L., Mancuso, M., 2001. Hydrodynamic and hydrochemical complexity of Naples urban area: some interpretations. Geologia Tecnica e Ambientale, 2, 35-54.
- Chandrasekharam, D., Bundschuh, J., 2002. Geochemistry of thermal waters and thermal gases, in: Chandrasekharam, D., Bundschuh, J., Geothermal energy resources for developing countries. Swets & Zeitlinger Publishers, Lisse, The Netherlands, 253-267.
- Chiodini, G., Caliro, S., Cardellini, C., Granieri, D., Avino, R., Baldini, A., Donnini, M., Minopoli, C., 2010. Long-term variations of the Campi Flegrei, Italy, volcanic system as revealed by the monitoring of hydrothermal activity. Journal of Geophysical Research: Solid Earth, 115(B3).
- Chiodini, G., Caliro, S., De Martino, P., Avino, R., Gherardi, F., 2012. Early signals of new volcanic unrest at Campi Flegrei caldera? Insights from geochemical data and physical simulations. Geology, 40(10), 943-946.
- Chiodini, G., Frondini, F., Cardellini, C., Granieri, D., Marini, L., Ventura, G., 2001. CO2 degassing and energy release at Solfatara volcano, Campi Flegrei, Italy. Journal of Geophysical Research: Solid Earth, 106(B8), 16213-16221.
- Chiodini, G., Todesco, M., Caliro, S., Del Gaudio, C., Macedonio, G., Russo, M., 2003. Magma degassing as a trigger of bradyseismic events: The case of Phlegrean Fields (Italy). Geophysical Research Letters, 30(8).
- Cortecci, G., Noto, P., Panichi, C., 1978. Environmental isotopic study of the Campi Flegrei (Naples, Italy) geothermal field. Journal of Hydrology, 36(1-2), 143-159.
- Dall'Aglio, M., Martini, M., Tonani, F., 1972. Rilevamento geochimico delle emanazioni vulcaniche dei Campi Flegrei, Quaderni De 'La Ricerca Scientifica'. CNR 83, 152–181.
- D'Auria, L., Pepe, S., Castaldo, R., Giudicepietro, F., Macedonio, G., Ricciolino, P., Tizzani, P., Casu, F., Lanari, R., Manzo, M., Martini, M., Sansosti, E., Zinno, I., 2015. Magma injection beneath the urban area of Naples: a new mechanism for the 2012–2013 volcanic unrest at Campi Flegrei caldera. Scientific reports, 5, 13100.
- De Vita, P., Allocca, V., Celico, F., Fabbrocino, S., Mattia, C., Monacelli, G., Musilli, I., Piscopo, V., Scalise, A. R., Summa, G., Tranfaglia, G., Celico, P., 2018. Hydrogeology of continental southern Italy, Journal of Maps, 14:2, 230-241, DOI: 10.1080/17445647.2018.1454352
- Del Gaudio, C., Aquino, I., Ricciardi, G. P., Ricco, C., Scandone, R., 2010. Unrest episodes at Campi Flegrei: A reconstruction of vertical ground movements during 1905–2009. Journal of Volcanology and Geothermal Research, 195(1), 48-56.
- De Vivo, B., Lima, A., 2008. Characterization and remediation of a brownfield site: the Bagnoli case in Italy. In De Vivo, B., Belkin, H. E., Lima, A. (Eds.), Environmental geochemistry. Site characterization. Data analysis and case histories. Elsevier, The Netherland: Amsterdam, pp. 356–384.
- Di Girolamo, P., 1978. Geotectonic settings of Miocene-Quaternary volcanism in and around the eastern Tyrrhenian Sea border (Italy) as deduced from major element geochemistry. Bulletin Volcanologique, 41(3), 229-250.
- Dominech, S., Yang, S., Ebrahimi, P., Aruta, A., Guarino, A., Gramazio, A., Albanese, S., 2021. A new approach to determine geochemical fingerprint of contaminants in stream sediments of southern Italy. Goldschmidt2021• Virtual• 4-9 July.
- Dotsika, E., 2015. H–O–C–S isotope and geochemical assessment of the geothermal area of Central Greece. Journal of Geochemical Exploration, 150, 1-15.
- Ebrahimi, P., Albanese, S., Esposito, L., Zuzolo, D., Cicchella, D., 2021. Coupling compositional data analysis (CoDA) with hierarchical cluster analysis (HCA) for preliminary understanding of the dynamics of a complex water distribution system: the Naples (South Italy) case study. Environmental Science: Water Research & Technology, 7(6), 1060-1077.
- Egozcue, J. J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Mathematical Geology, 37(7), 795-828.
- Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barcelo-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Mathematical Geology, 35(3), 279-300.

- Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis, Springer International Publishing, 10.1007/978-3-319-96422-5.
- Ghiara, M. R., Stanzione, D., Bellucci, L., Panichi, C., 1988. Preliminary observations on the chemical variations of the solutes monitored in the thermal waters of the Phlegraean Fields in the period 1983–1985. Rendiconti della Societa Italiana di Mineralogia e Petrologia, 43, 1139-1150.
- Giggenbach, W. F., 1991. Chemical techniques in geothermal exploration. In D'Amore, F. (co-ordinator), Application of Geochemistry in Geothermal Reservoir Development. UNITAR, 119-144.
- Guglielminetti, M., 1986. Mofete geothermal field. Geothermics, 15(5-6), 781-790.
- Giudicepietro, F., Chiodini, G., Caliro, S., De Cesare, W., Esposito, A. M., Galluzzo, D., Lo Bascio, D., Macedonio, G., Orazi, M., Ricciolino, P., Vandemeulebrouck, J., 2019. Insight into Campi Flegrei caldera unrest through seismic tremor measurements at Pisciarelli fumarolic field. Geochemistry, Geophysics, Geosystems, 20(11), 5544-5555.
- ISPRA (Istituto superiore per la protezione e la ricerca ambientale), 2018. Carta Geologica d'Italia. 1:50,000 Sheets: 446-447 Napoli.
- Keshavarzi, B., Ebrahimi, P., Moore, F., 2015. A GIS-based approach for detecting pollution sources and bioavailability of metals in coastal and marine sediments of Chabahar Bay, SE Iran. Geochemistry, 75(2), 185-195. https://doi.org/10.1016/j.chemer.2014.11.003
- Kynčlová, P., Hron, K., Filzmoser, P., 2017. Correlation between compositional parts based on symmetric balances. Mathematical Geosciences, 49(6), 777-796.
- McQuitty, L. L., 1966. Similarity analysis by reciprocal pairs for discrete and continuous data. Educational and Psychological Measurement, 26(4), 825-831. doi: 10.1177/001316446602600402.
- Moeck, I. S., 2014. Catalog of geothermal play types based on geologic controls. Renewable and Sustainable Energy Reviews, 37, 867-882.
- Pearson, K., 1897. Mathematical contributions to the theory of evolution.—on a form of spurious correlation which may arise when indices are used in the measurement of organs. Proceedings of the Royal Society of London, 60 (359-367), 489-498.
- Peccerillo, A., 1985. Roman comagmatic province (central Italy): evidence for subduction-related magma genesis. Geology, 13(2), 103-106.
- Piper, A. M., 1944. A graphic procedure in the geochemical interpretation of water-analyses. Eos, Transactions American Geophysical Union, 25(6), 914-928.
- Piochi, M., Kilburn, C. R. J., Di Vito, M. A., Mormone, A., Tramelli, A., Troise, C., De Natale, G., 2014. The volcanic and geothermally active Campi Flegrei caldera: an integrated multidisciplinary image of its buried structure. International Journal of Earth Sciences, 103(2), 401-421.
- R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
- Reimann, C., Filzmoser, P., Garrett, R., Dutter, R., 2008. Statistical data analysis explained: applied environmental statistics with R. John Wiley & Sons. West Sussex, England.
- Reimann, C., Filzmoser, P., Hron, K., Kynčlová, P., Garrett, R. G., 2017. A new method for correlation analysis of compositional (environmental) data–a worked example. Science of the Total Environment, 607, 965-971.
- Rosi, M., Sbrana, A., 1987. Phlegraean Fields, Quaderni de "La ricerca scientifica", Consiglio Nazionale delle Ricerche, 9, 175 pp.
- Sellerino, M., Forte, G., Ducci, D., 2019. Identification of the natural background levels in the Phlaegrean fields groundwater body (Southern Italy). Journal of Geochemical Exploration, 200, 181-192.
- Scarpati, C., Perrotta, A., Lepore, S., Calvert, A., 2013. Eruptive history of Neapolitan volcanoes: constraints from 40Ar–39Ar dating. Geological Magazine, 150(3), 412-425.
- Shelton, J. L., Engle, M. A., Buccianti, A., Blondes, M. S., 2018. The isometric log-ratio (ilr)-ion plot: A proposed alternative to the Piper diagram. Journal of Geochemical Exploration, 190, 130-141.
- Somma, R., Ebrahimi, P., Troise, C., De Natale, G., Guarino, A., Cicchella, D., Albanese, S., 2021. The first application of compositional data analysis (CoDA) in a multivariate perspective for detection of pollution source in sea sediments: The Pozzuoli Bay (Italy) case study. Chemosphere, 274, 129955.
- Stellato, L., Coda, S., Arienzo, M., De Vita, P., Di Rienzo, B., D'Onofrio, A., Ferrara, L., Marzaioli, F., Trifuoggi, M., Allocca, V., 2020. Natural and Anthropogenic Groundwater Contamination in a Coastal

Volcanic-Sedimentary Aquifer: The Case of the Archaeological Site of Cumae (Phlegraean Fields, Southern Italy). Water, 12(12), 3463.

- Trasatti, E., Polcari, M., Bonafede, M., Stramondo, S., 2015. Geodetic constraints to the source mechanism of the 2011–2013 unrest at Campi Flegrei (Italy) caldera. Geophysical Research Letters, 42(10), 3847-3854.
- Valentino, G. M., Cortecci, G., Franco, E., Stanzione, D., 1999. Chemical and isotopic compositions of minerals and waters from the Campi Flegrei volcanic system, Naples, Italy. Journal of Volcanology and Geothermal Research, 91(2-4), 329-344.
- Valentino, G. M., Stanzione, D., 2003. Source processes of the thermal waters from the Phlegraean Fields (Naples, Italy) by means of the study of selected minor and trace elements distribution. Chemical Geology, 194(4), 245-274.
- Valentino, G. M., Stanzione, D., 2004. Geochemical monitoring of the thermal waters of the Phlegraean Fields. Journal of Volcanology and Geothermal Research, 133(1-4), 261-289.
- Van den Boogaart, K. G., Tolosana-Delgado, R., 2013. Analyzing compositional data with R (Vol. 122). Berlin: Springer, 10.1007/978-3-642-36809-7.
- Vitale, S., Isaia, R., 2014. Fractures and faults in volcanic rocks (Campi Flegrei, southern Italy): insight into volcano-tectonic processes. International Journal of Earth Sciences, 103(3), 801-819.
- Washington, H. S., 1906. ‘‘The Roman Comagmatic Region.’’ Carnegie Inst. of Washington. Vol. 57, p. 199.

679 **Table 1** The measured values of ions, elements and physico-chemical parameters in Campi Flegrei groundwater samples

ID	Water type	Dominant anion group [a]	HCA group [b]	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	Cl^{-}	SO_4^{2-}	HCO_3^{-}	NO_3^{-}	F^{-}	Li	As	B	EC	T	pH
W1	Na–SO_4	SGW	$C4_{SGW}$	38.3	0.4	768.0	184.0	126.0	1040.0	817.0	0.7	9.6	423.0	168.1	1457.5	3400	86.2	7.0
W2	Na–SO_4	SGW	$C4_{SGW}$	89.2	2.0	1000.0	288.0	119.0	1880.0	987.0	<D.L.	3.1	629.0	1526.3	27900.0	4480	79.4	6.8
W3	Ca–Cl	ClGW	$C2_{MGW}$	131.0	13.9	135.0	53.4	214.0	223.0	216.0	138.0	1.2	<D.L.	14.4	314.4	1420	16.4	6.9
W4	Na–Cl	ClGW	$C2_{MGW}$	34.9	7.3	482.0	133.0	413.0	205.0	492.0	101.0	7.3	290.0	460.1	1040.0	2550	36	7.7
W5	Na–Cl	ClGW	$C3_{ClGW}$	164.0	11.4	7740.0	176.0	10900.0	1410.0	2260.0	<D.L.	1.3	11000.0	3.2	88812.5	23000	20	6.2
W6	Na–Cl	ClGW	$C2_{MGW}$	65.6	10.4	261.0	53.2	316.0	115.0	374.0	45.9	4.4	149.0	168.9	938.4	1690	21.3	7.1
W7	Na–Cl	ClGW	$C2_{MGW}$	37.1	3.1	542.0	113.0	511.0	204.0	474.0	104.0	7.7	275.0	419.5	5378.8	2650	44.3	7.3
W8	Na–Cl	ClGW	$C2_{MGW}$	263.0	57.0	2020.0	327.0	2880.0	376.0	1300.0	2.8	2.1	1230.0	322.5	13837.5	8110	56	6.6
W9	Na–Cl	ClGW	$C2_{MGW}$	82.1	2.9	495.0	115.0	417.0	254.0	478.0	129.0	6.0	312.0	506.9	1656.3	2480	35.9	7.8
W10	Na–Cl	ClGW	$C3_{ClGW}$	564.0	872.0	11200.0	761.0	20700.0	2240.0	807.0	7.9	2.4	4410.0	3538.8	23800.0	37000	67	6.1
W11	Na–Cl	ClGW	$C3_{ClGW}$	382.0	244.0	7870.0	539.0	14800.0	317.0	48.8	1.2	6.0	4210.0	<D.L.	9927.5	33100	43.7	6.5
W12	Na–Cl	ClGW	$C3_{ClGW}$	182.0	12.8	3850.0	250.0	6330.0	200.0	810.0	65.6	2.7	6660.0	1610.0	10470.0	14780	29.5	6.5
W13	Na–Cl	ClGW	$C2_{MGW}$	59.6	0.7	1070.0	378.0	932.0	1090.0	758.0	0.7	3.8	645.0	968.9	19825.0	5160	75	7.3
W14	Na–Cl	ClGW	$C3_{ClGW}$	85.0	9.8	108.0	20.6	349.0	7.2	83.0	4.3	0.4	<D.L.	11.9	191.3	795	18.3	7.2
W15	Ca–Cl	ClGW	$C2_{MGW}$	209.0	4.7	212.0	30.2	234.0	267.0	262.0	174.0	4.5	<D.L.	333.5	2888.8	17160	36.4	7.5
W16	Na–Cl	ClGW	$C2_{MGW}$	21.0	0.9	232.0	37.3	234.0	27.6	299.0	<D.L.	12.7	<D.L.	158.9	1407.5	1230		7.6
W17	Na–Cl	ClGW	$C2_{MGW}$	63.0	12.0	559.0	493.0	916.0	343.0	580.0	1.1	9.5	<D.L.	118.6	3170.0	2310	26.4	5.7
W18	Na–Cl	ClGW	$C2_{MGW}$	424.0	31.4	1690.0	282.0	2230.0	368.0	1190.0	1.7	1.8	329.0	524.6	7246.3	5600	21.4	6.4
W19	Na–Cl	ClGW	$C3_{ClGW}$	516.0	526.0	7110.0	829.0	12600.0	1560.0	1620.0	1.7	3.0	2610.0	23.2	11098.8	25500	18.7	6.0
W20	Na–Cl	ClGW	$C2_{MGW}$	99.0	8.4	503.0	95.9	586.0	139.0	586.0	12.8	5.8	295.0	258.3	5226.3	2680	23.5	7.6
W21	Na–Cl	ClGW	$C2_{MGW}$	127.0	30.0	843.0	143.0	931.0	134.0	1150.0	16.0	2.9	264.0	473.8	7005.0	3380	36.2	6.9
W22	Na–Cl	ClGW	$C3_{ClGW}$	309.0	60.5	7370.0	207.0	11000.0	611.0	371.0	132.0	13.7	11400.0	6875.0	31700.0	26000	77	6.9
W23	Na–Cl	ClGW	$C3_{ClGW}$	168.0	77.9	2430.0	208.0	3990.0	405.0	870.0	33.0	6.0	2380.0	793.6	13975.0	11490	32	7.0
W24	Na–Cl	ClGW	$C3_{ClGW}$	188.0	29.0	2070.0	148.0	3560.0	194.0	336.0	80.9	2.4	2330.0	262.9	6111.3	9400	27.3	7.1
W25	Na–HCO_3	BGW	$C1_{BGW}$	58.8	7.9	212.0	69.3	100.0	129.0	309.0	65.3	4.9	<D.L.	41.5	595.1	1021	21	7.4
W26	Na–HCO_3	BGW	$C1_{BGW}$	317.0	26.1	498.0	138.0	510.0	613.0	1000.0	0.4	0.4	363.0	724.4	16587.5	2770	40.5	6.2
W27	Na–HCO_3	BGW	$C1_{BGW}$	180.0	68.4	637.0	99.4	551.0	430.0	1130.0	355.0	3.1	196.0	33.6	1058.9	3770	22.7	7.1
W28	Na–HCO_3	BGW	$C1_{BGW}$	99.9	9.3	116.0	50.8	88.9	195.0	265.0	101.0	4.1	<D.L.	34.5	425.7	787	17.5	7.4
W29	Na–HCO_3	BGW	$C1_{BGW}$	42.0	5.8	246.0	102.0	153.0	89.1	614.0	73.4	1.3	<D.L.	199.6	334.1	1550	25.1	7.2

W30	$Na–HCO_3$	BGW	$C1_{BGW}$	127.0	7.3	147.0	75.0	129.0	228.0	295.0	168.0	6.7	<D.L.	44.2	252.1	1360	17.4	7.4
W31	$Na–HCO_3$	BGW	$C1_{BGW}$	162.0	17.4	724.0	134.0	179.0	256.0	1870.0	0.6	3.1	883.0	154.8	1776.3	2490	26	6.8
W32	$Na–HCO_3$	BGW	$C1_{BGW}$	31.5	1.9	263.0	35.1	162.0	67.8	447.0	50.1	9.2	<D.L.	73.5	781.0	1340	22.3	7.6
W33	$Na–HCO_3$	BGW	$C1_{BGW}$	118.0	29.9	823.0	110.0	412.0	247.0	1700.0	7.7	2.2	980.0	292.4	4681.3	3200	35.7	6.3
W34	$Na–HCO_3$	BGW	$C1_{BGW}$	149.0	26.7	764.0	136.0	462.0	148.0	1590.0	0.6	3.2	770.0	166.4	5713.8	3060	33.3	6.4
W35	$Na–HCO_3$	BGW	$C1_{BGW}$	3.7	0.1	299.0	36.5	174.0	7.9	509.0	0.4	12.3	<D.L.	234.6	2432.5	1218	32.8	8.6
W36	$Na–HCO_3$	BGW	$C1_{BGW}$	57.1	2.3	90.1	54.8	69.5	127.0	238.0	10.3	7.5	<D.L.	28.1	943.4	633	16.7	7.8
W37	$Na–HCO_3$	BGW	$C1_{BGW}$	32.2	2.3	43.8	41.4	28.3	5.1	204.0	0.8	1.9	<D.L.	6.9	1666.0	350	19.1	8.1
W38	$Ca–HCO_3$	BGW	$C1_{BGW}$	215.0	9.0	132.0	76.4	140.0	254.0	512.0	176.0	3.6	<D.L.	72.3	964.3	1159	18.8	7.3
W39	$Na–HCO_3$	BGW	$C2_{MGW}$	39.4	5.1	1690.0	157.0	1100.0	530.0	1940.0	24.1	3.7	928.0	1590.0	13437.5	6340	53.4	6.9
W40	$Na–HCO_3$	BGW	$C1_{BGW}$	107.0	17.9	400.0	100.0	259.0	205.0	926.0	0.8	4.8	368.0	140.6	4825.0	1878	23.6	6.2
W41	$Ca–HCO_3$	BGW	$C1_{BGW}$	83.2	6.6	88.3	27.4	53.7	33.6	352.0	37.9	2.3	<D.L.	132.4	204.1	638	22.4	6.7
W42	$Ca–HCO_3$	BGW	$C1_{BGW}$	120.0	10.0	130.0	32.8	86.5	41.9	509.0	34.5	3.8	<D.L.	108.7	364.6	786	21.8	6.7
W43	$Na–HCO_3$	BGW	$C1_{BGW}$	66.0	2.9	909.0	236.0	202.0	696.0	1510.0	62.6	8.8	386.0	437.0	9186.3	3880	47.7	7.0
W44	$Na–HCO_3$	BGW	$C1_{BGW}$	19.0	0.4	213.0	27.9	114.0	114.0	244.0	84.1	12.5	<D.L.	79.8	1838.8	1020	21.8	7.4
Min				3.7	0.1	43.8	20.6	28.3	5.1	48.8	<D.L.	0.4	<D.L.	<D.L.	191.3	350	16.4	5.7
Max				564.0	872.0	11200.0	829.0	20700.0	2240.0	2260.0	355.0	13.7	11400.0	6875.0	88812.5	37000	86.2	8.6
Average				143.9	51.9	1567.8	172.8	2278.7	409.7	757.6	56.3	5.0	2026.5	561.3	8260.1	6469	34.1	7.0
CVR				89.9	115.5	108.1	86.6	100.5	83.5	76.7	145.1	75.4	148.2	125.7	130.9	87.6	44.4	8.4
GEW (Mofete 2) [c]				480	1	5090	1180	10200	3160	41	-	-	13000	11000	140000	-	337	6.0
GEW (Mofete 1) [c]				555	-	10025	1230	17710	615	81	-	-	25000	9000	125000	-	250	7.5
GEW (Mofete 1) [c]				1281	5	12589	2342	22810	670	46	-	-	28000	11000	110000	-	250	6.5
LSW [d]				447	1397	12250	446	22390	2906	161	-	-	170	2	4410	-	20	8.1

680 <D.L.: below detection limit; CVR=(MAD/Median)×100 where MAD is median absolute deviation (Reimann et al., 2008); GEW: geothermal water; LSW: local seawater. All values are in mg/l except for
681 Li, As and B (μg/l), electric conductivity (EC; μS/cm), temperature (T; °C) and pH (unitless). [a] The groups are based on the dominant anion species in meq/l: SGW, ClGW and BGW stand for sulfate–,
682 chlorine– and bicarbonate–rich groundwaters, respectively; [b] The groups are based on hierarchical cluster analysis: $C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$ denotes the bicarbonate–rich, mixed, chlorine–rich
683 and sulfate–rich groundwaters, respectively; [c] Guglielminetti (1986); [d] Aiuppa et al. (2006)

684

685

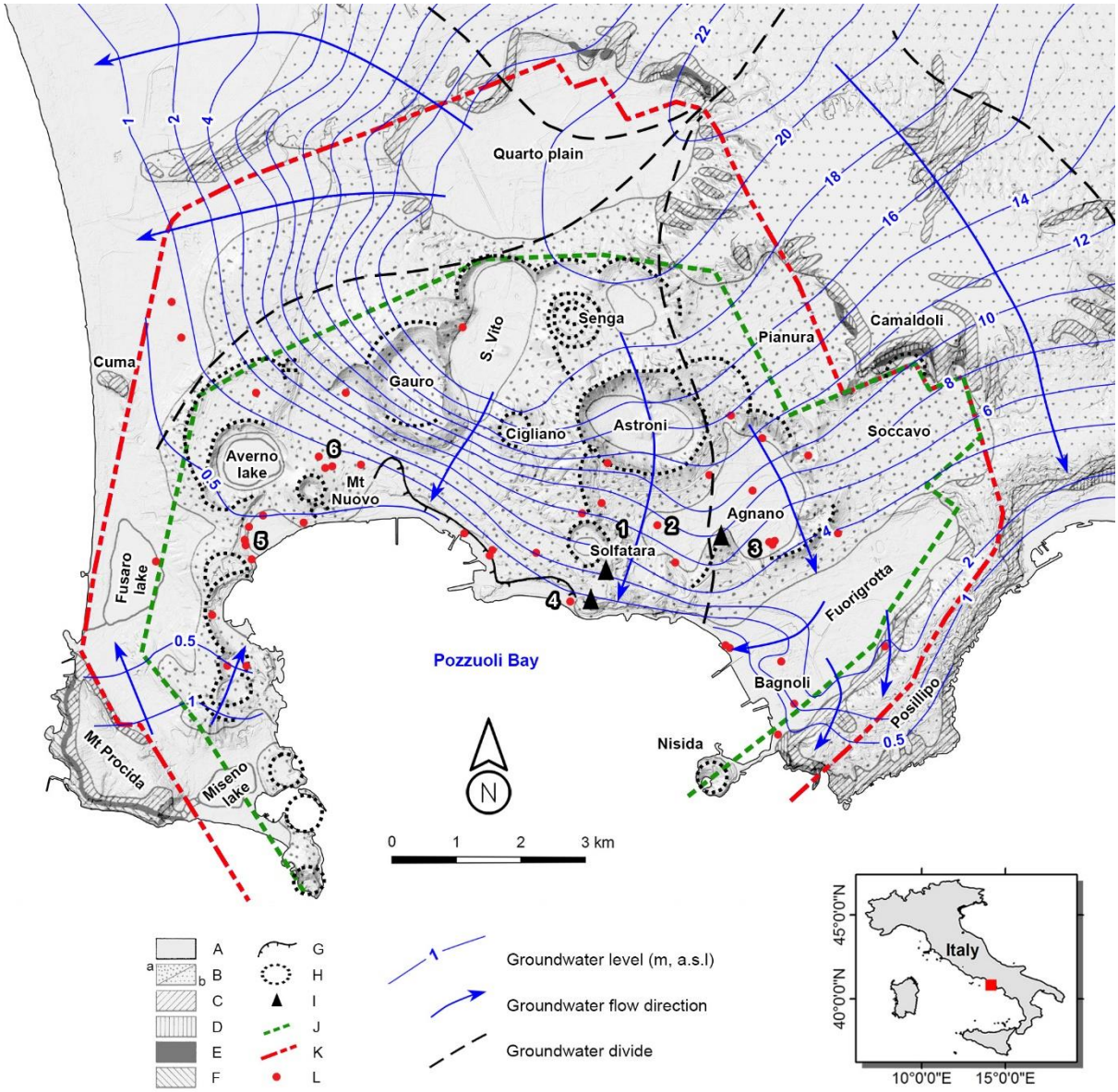


Fig. 1 The simplified geological map of Campi Flegrei (after Valentino et al., 1999). The groundwater level, groundwater flow direction, groundwater divide (after De Vita et al., 2018) together with location of (1) Pisciarelli, (2) Hotel Tennis, (3) Agnano Terme, (4) Terme Puteolane, (5) Stufe di Nerone and (6) Damiani are also represented. Legend: (A) Recent active continental sediments; (B) Volcanics younger than 12 ka: a proximal, mostly pyroclastic-flow and surge deposits, b distal, mostly fallout deposits; (C) Neapolitan Yellow Tuff (NYT; 12 ka B.P.); (D) Volcanics erupted 35-12 ka B.P.; (E) Campanian Ignimbrite (CI; 35 ka B.P.); (F) Volcanics older than 35 ka; (G) Edge of La Starza marine cliff; (H) Crater rims of volcanoes younger than 12 ka (ISPRA, 2018); (I) Lava domes; (J) NYT caldera rim (Vitale and Isaia, 2014); (K) CI caldera rim (Vitale and Isaia, 2014); (L) sampling points.

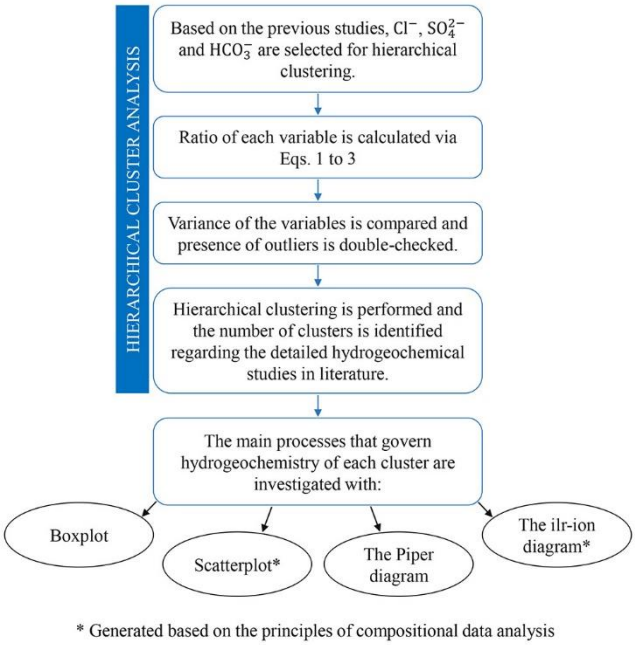


Fig. 2 The flowchart demonstrating the main data treatment and visualization steps.

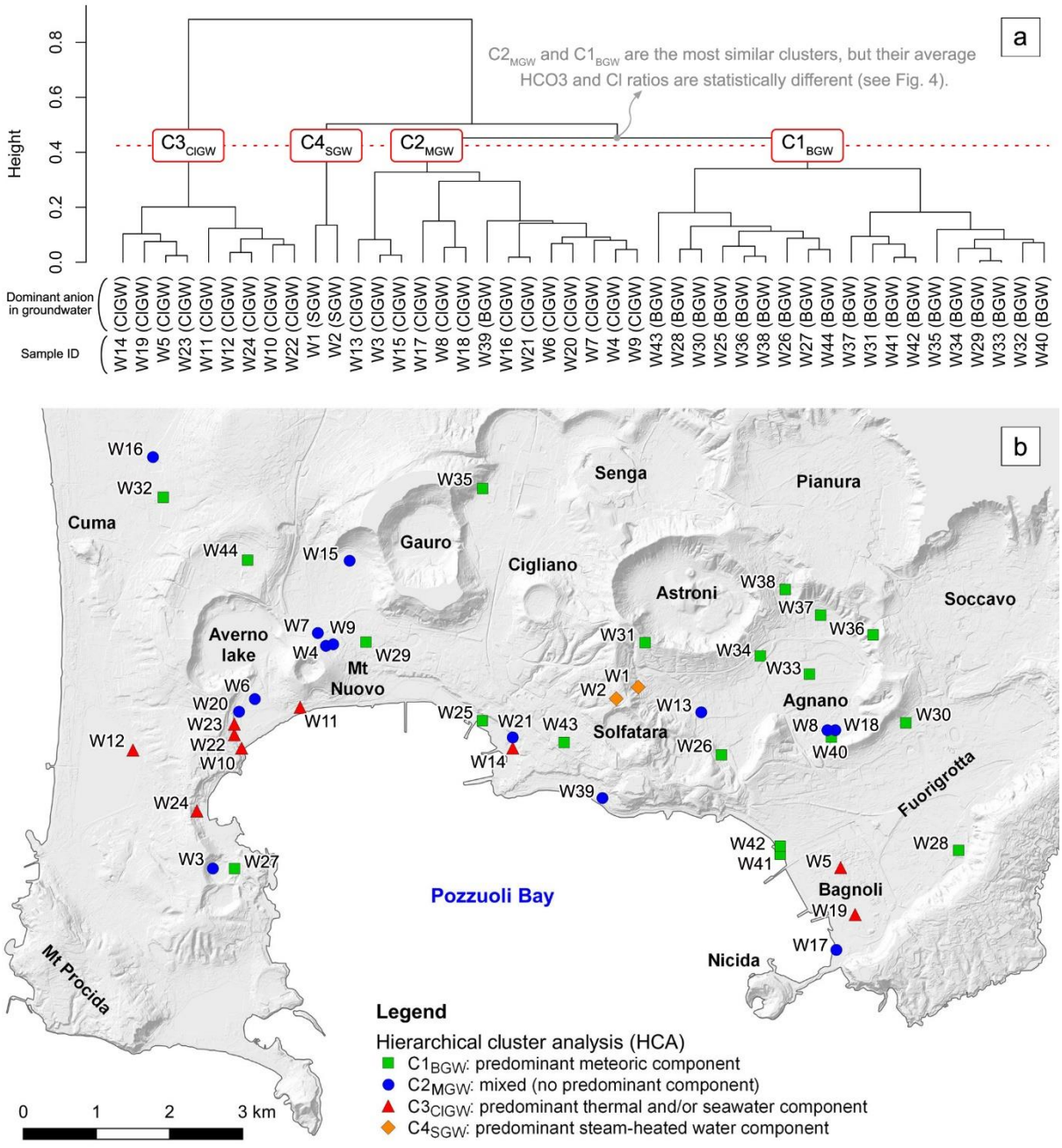


Fig. 3 a The dendrogram of hierarchical cluster analysis; and **b** spatial distribution of each cluster. The height on the vertical axis of the dendrogram states the distance measure, indicating higher similarity at a lower value.

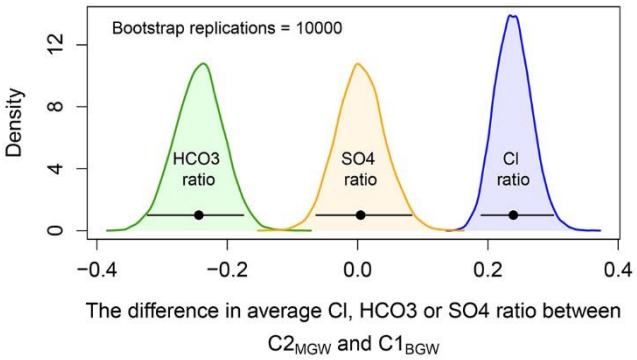


Fig. 4 Density plot of the difference in average HCO_3^-, SO_4^{2-} or Cl^- ratio between the mixed ($C2_{MGW}$) and bicarbonate–rich ($C1_{BGW}$) groundwaters. The dots show the values obtained from the dataset in the present article and the horizontal line indicates the uncertainty based on the 95% confidence interval of 10000 replications. In each bootstrap replication, for instance, average HCO_3^- ratio in the bicarbonate–rich groundwater is subtracted from that of the mixed groundwater. Thus, if the middle 95% of the distribution is different from zero, the difference is statistically significant.

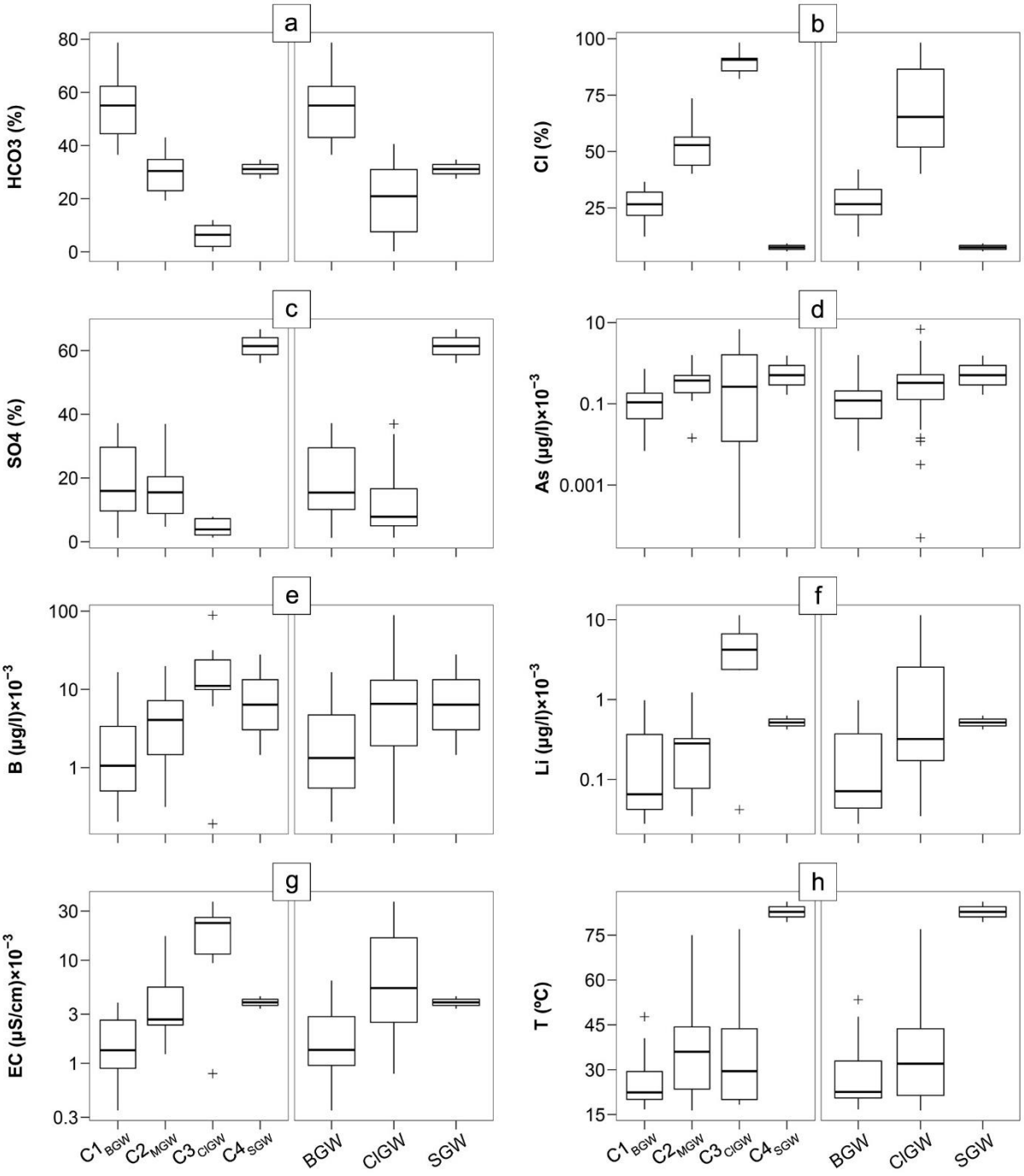


Fig. 5 Variation of selected parameters in the HCA groups and dominant anion groups (the left and right panels in each paired graph, respectively). To generate Figs. 5a to c, the following calculations were made using bicarbonate, chloride and sulfate in meq/l: $\mathrm{HCO_3^-} = (\mathrm{HCO_3^-}/(\mathrm{HCO_3^-} + \mathrm{Cl^-} + \mathrm{SO_4^{2-}})) \times 100$; $\mathrm{Cl^-} = (\mathrm{Cl^-}/(\mathrm{HCO_3^-} + \mathrm{Cl^-} + \mathrm{SO_4^{2-}})) \times 100$; and $\mathrm{SO_4^{2-}} = (\mathrm{SO_4^{2-}}/(\mathrm{HCO_3^-} + \mathrm{Cl^-} + \mathrm{SO_4^{2-}})) \times 100$. Scale of y-axis is logarithmic in As, B, Li and EC (**d** to **g**, respectively) boxplots. EC stands for electric conductivity.

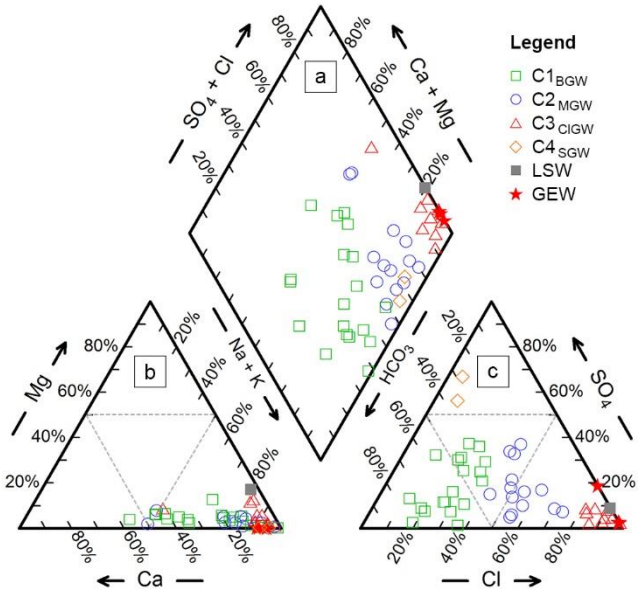


Fig. 6 Piper diagram showing composition of Campi Flegrei groundwaters, grouped according to HCA results, together with geothermal water (GEW) and local seawater (LSW): **a** the diamond-shaped field; **b** the cation field; and **c** the anion field.

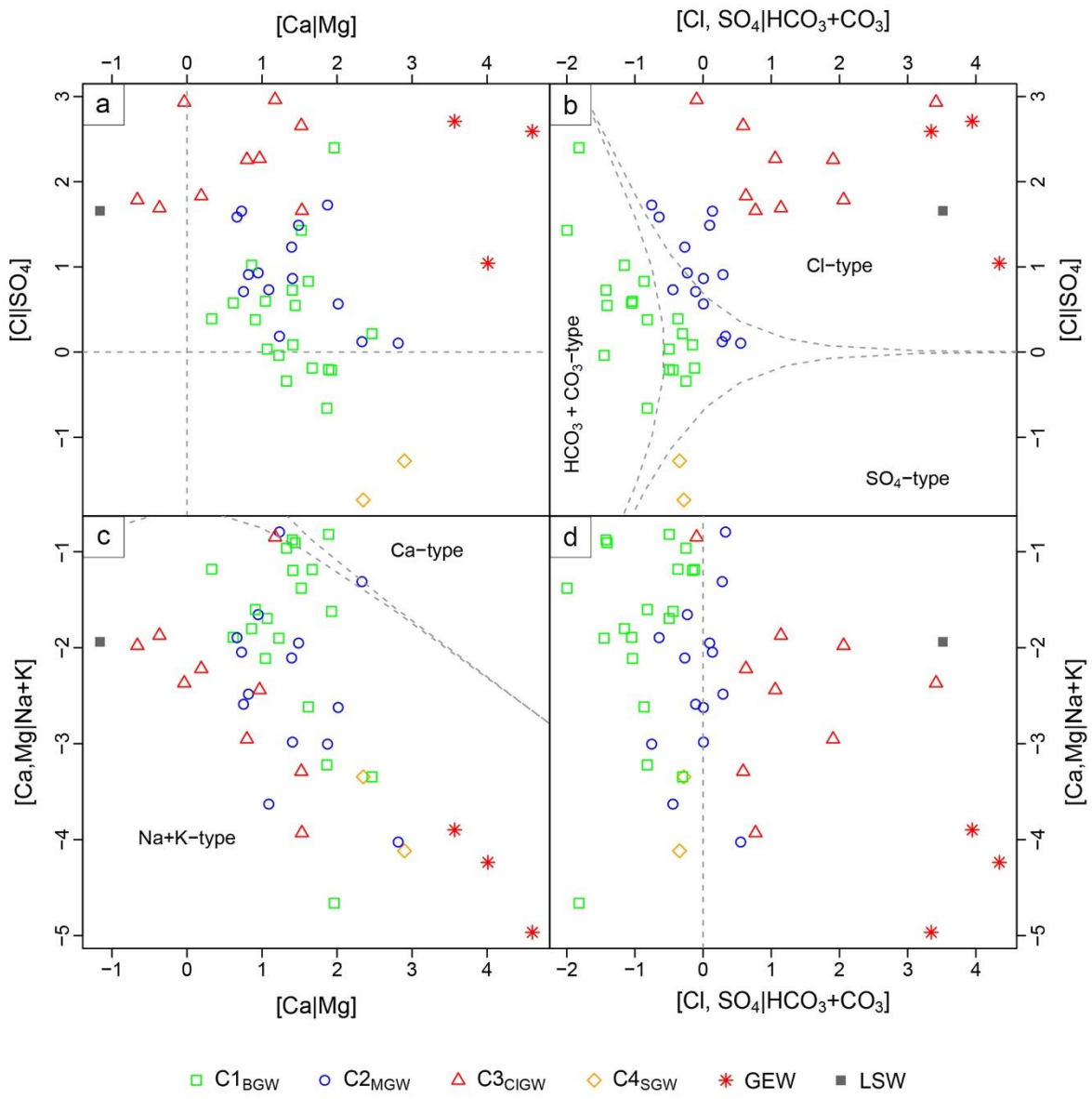


Fig. 7 Chemical composition of the collected groundwater samples on the ilr-ion plot: **a** the information in this panel is not specifically presented in the Piper diagram; **b** corresponds to the anion field of the Piper diagram; **c** corresponds to the cation field of the Piper diagram; and **d** corresponds to the diamond-shaped field of the Piper diagram.

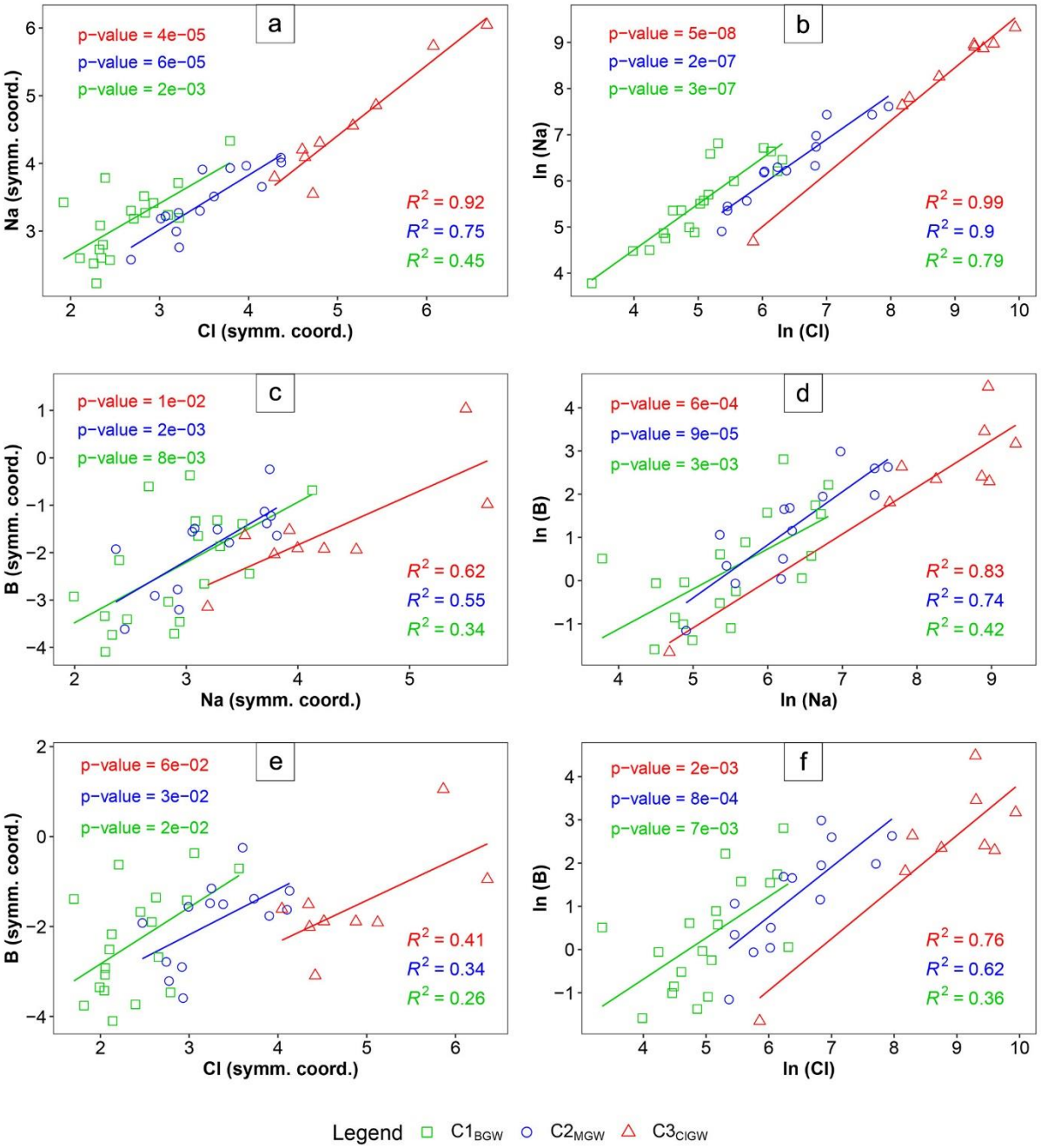


Fig. 8 The relationship between Na, Cl and B considering the symmetric coordinates (**a**, **c** and **e**) and the classical log-transformed data (**b**, **d** and **f**). The ions (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, HCO_3^-, SO_4^{2-}, Cl^-, NO_3^- and F^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity were used for calculation of the symmetric coordinates. The sulfate–rich groundwaters are excluded because of few samples.

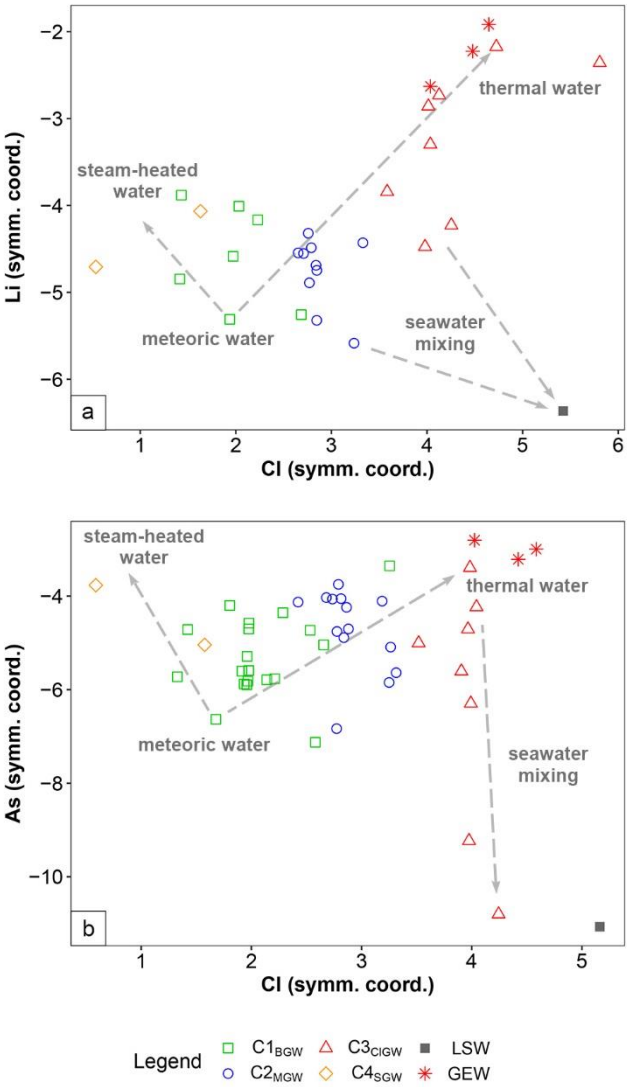


Fig. 9 Li–Cl^- and As–Cl^- scatterplots (**a** and **b**, respectively) generated by the symmetric coordinates. The ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-} and Cl^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity were used for calculation of the symmetric coordinates. The samples with a lithium content of below detection limit are excluded in the Li–Cl^- panel.

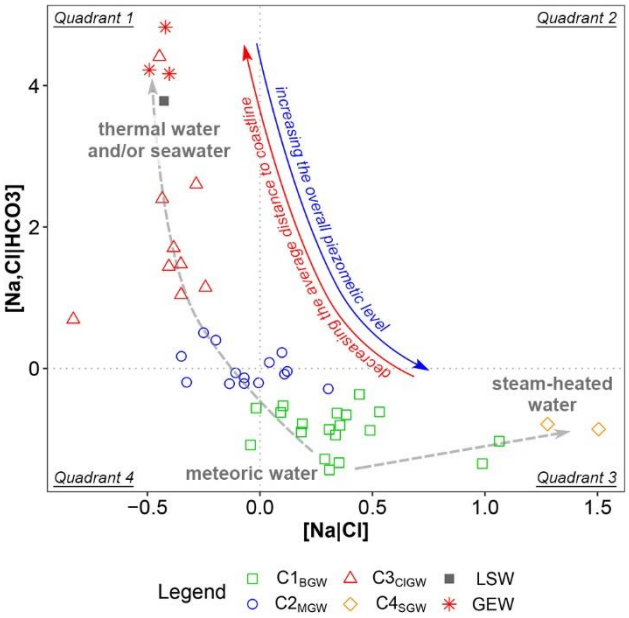


Fig. 10 The relationship between the isometric log-ratios computed using Na^+, Cl^- and HCO_3^-.

Supplementary material for

The main hydrogeochemical processes in Campi Flegrei volcanic aquifer, south Italy: Proposing anion ratios for hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry

Pooria Ebrahimi [a]*, Annalise Guarino [a], Vincenzo Allocca [a], Stefano Caliro [b], Rosario Avino [b], Emanuela Bagnato [b], Francesco Capecchiacci [b,c], Antonio Carandente [b], Carmine Minopoli [b], Alessandro Santi [b], Stefano Albanese [a]

[a] Department of Earth, Environmental and Resources Sciences, University of Naples Federico II, Naples 80126, Italy
[b] Istituto Nazionale di Geofisica e Vulcanologia sez. "Osservatorio Vesuviano", Naples 80124, Italy
[c] Department of Earth Sciences, University of Florence, Florence 50121, Italy

* Corresponding author: pooria.ebrahimi@unina.it; pooria.ebrahimi@gmail.com

Table S1 Descriptive ID and coordinates (UTM zone 33N) of the collected samples in May 2019.

Table S2 The shortest distance (m) of the groundwater samples from the coastline. $C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$ refer to the bicarbonate–rich, mixed, chlorine–rich and sulfate–rich groundwaters, respectively.

Table S3 Literature data on major anion levels (mg/l) in the groundwater samples from Puteolane well (Valentino and Stanzione, 2004). The dominant anion species is chloride in all samples.

Fig. S1 Boxplots of major anions (**a** to **c**) and their ratios (**d** to **f**) calculated via equations 1 to 3 in the article. Variance of each variable is also represented above the corresponding chart.

Table S1 Descriptive ID and coordinates (UTM zone 33N) of the collected samples in May 2019.

ID	Descriptive ID	Sampling date	X	Y	ID	Descriptive ID	Sampling date	X	Y
W1	CF39 Agriturismo	14.5.2019	427611	4520624	**W23**	Stufe di Nerone spring	6.5.2019	422083	4520083
W2	Hotel Gli Dei	14.5.2019	427310	4520465	**W24**	Terme di Baia	6.5.2019	421576	4518931
W3	Agriturismo fondi di Baia	6.5.2019	421795	4518144	**W25**	Sud Cantieri Navali	14.5.2019	425479	4520164
W4	Centro ippico	7.5.2019	423341	4521188	**W26**	Carannante	8.5.2019	428748	4519697
W5	Pz Bagnoli CFDDP	6.5.2019	430379	4518158	**W27**	Castello di Baia	6.5.2019	422090	4518143
W6	Costagliola	6.5.2019	422368	4520460	**W28**	Cataldo	10.5.2019	431993	4518392
W7	Damiani	7.5.2019	423230	4521363	**W29**	Conte	9.5.2019	423888	4521240
W8	De Pisis	7.5.2019	430254	4519984	**W30**	Di Napoli	9.5.2019	431270	4520135
W9	Esposito	7.5.2019	423440	4521209	**W31**	Frolla	10.5.2019	427705	4521232
W10	Franco allo Scoglio	6.5.2019	422187	4519790	**W32**	Guardascione	8.5.2019	421116	4523218
W11	Giardinetto Via Miliscola	9.5.2019	422988	4520348	**W33**	Ippodromo centro pista	7.5.2019	429953	4520801
W12	Grotta dell'acqua	6.5.2019	420701	4519764	**W34**	Macars	8.5.2019	429281	4521050
W13	Hotel Tennis	7.5.2019	428473	4520279	**W35**	Marina Militare II	16.5.2019	425482	4523338
W14	Lopez	8.5.2019	425914	4519904	**W36**	Monte San Angelo (Di Rella)	10.5.2019	430822	4521338
W15	Parco Delta	7.5.2019	423664	4522351	**W37**	Parziale	7.5.2019	430107	4521608
W16	Parco Enea	7.5.2019	420975	4523768	**W38**	Pozzo di Dio	8.5.2019	429621	4521958
W17	Piezometro Coroglio	6.5.2019	430321	4517031	**W39**	Puteolane	8.5.2019	427122	4519107
W18	Pozzanghera Agnano Terme	7.5.2019	430283	4520019	**W40**	Agnano spring	7.5.2019	430220	4520007
W19	Pz Italsider	6.5.2019	430577	4517515	**W41**	Spiaggia Bagnoli	7.5.2019	429574	4518380
W20	Samuele	8.5.2019	422152	4520287	**W42**	Spiaggia Bagnoli 2	7.5.2019	429531	4518410
W21	Serapeo	6.5.2019	425876	4519823	**W43**	Tortorelli	8.5.2019	426598	4519866
W22	Stufe di Nerone well	6.5.2019	422092	4520008	**W44**	Urzo	7.5.2019	422270	4522360

Table S2 The shortest distance (m) of the groundwater samples from the coastline. $C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$ refer to the bicarbonate–rich, mixed, chlorine–rich and sulfate–rich groundwaters, respectively.

Cluster	ID	Distance to coast (m)
$C1_{BGW}$	W25	52
	W26	825
	W27	326
	W28	1168
	W29	864
	W30	2434
	W31	2243
	W32	1477
	W33	2203
	W34	2208
	W35	3001
	W36	3101
	W37	2996
	W38	3168
	W40	1654
	W41	40
	W42	40
	W43	671
	W44	2217
Average		1615
$C2_{MGW}$	W3	478
	W4	838
	W6	467
	W7	1021
	W8	1653
	W9	852
	W13	1380
	W15	1977
	W16	1397
	W17	48
	W18	1700
	W20	445
	W21	160
	W39	91
Average		893
$C3_{ClGW}$	W5	581
	W10	12
	W11	85
	W12	1023
	W14	246
	W19	360
	W22	256
	W23	321
	W24	150
Average		337
$C4_{SGW}$	W1	1642
	W2	1384
Average		1513

Table S3 Literature data on major anion levels (mg/l) in the groundwater samples from Puteolane well (Valentino and Stanzione, 2004). The dominant anion species is chloride in all samples.

Date	Cl^-	HCO_3^-	SO_4^{2-}
1985 [a]	2102	2599	624
1989 [a]	1650	2520	600
Feb–90	3425	2019	594
Feb–92	3886	1930	396
Sep–92	3891	1870	388
Feb–93	3896	1940	1152
May–93	3900	1933	1140
Sep–93	4200	1977	1197
Nov–93	4661	1874	1220
Jan–94	3830	2000	1000
Mar–94	5200	1919	1302
May–94	3680	1919	1038
Jul–94	4000	1950	1000
Sep–94	3540	2115	931
Nov–94	3450	2141	701
May–97	3620	2074	875
Jun–97	3988	1941	782
Jul–97	4363	1963	828
Sep–97	4240	2073	880
Nov–97	4114	1993	749
Mar–98	2662	2042	730
Jun–98	4845	2036	814
Sep–98	4700	2045	905
Jan–99	5115	1920	820
Feb–99	4897	2012	736
Mar–99	5210	1989	766
Apr–99	4800	2054	670
Jul–99	3600	2006	480
Sep–99	5120	2024	746

[a] Celico et al. (1992)

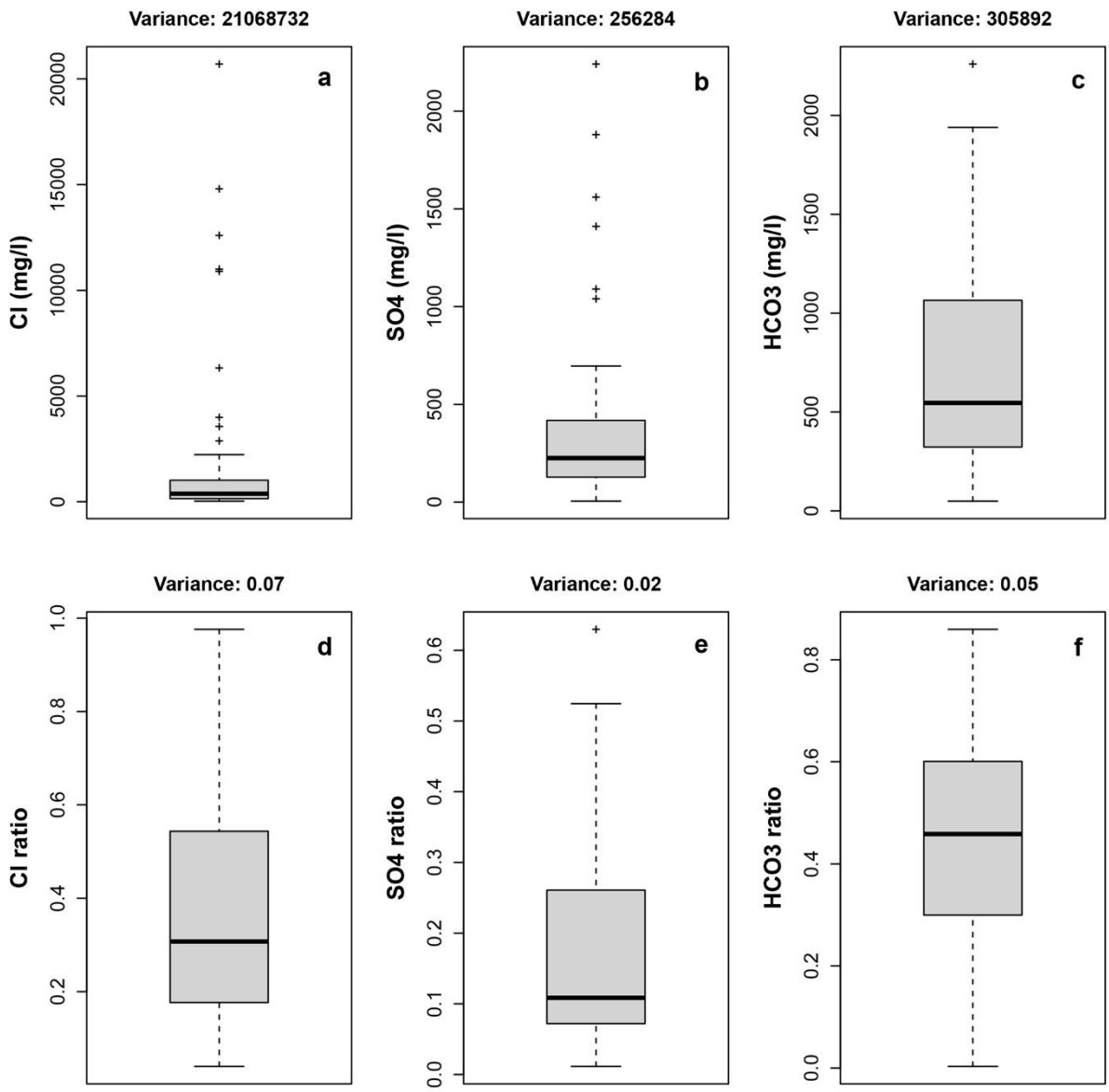


Fig. S1 Boxplots of major anions (**a** to **c**) and their ratios (**d** to **f**) calculated via equations 1 to 3 in the article. Variance of each variable is also represented above the corresponding chart.

References

- Celico, P., De Vita, P., Nikzad, F., Stanzione, D., & Vallario, A. (1992). Schema idrogeologico ed idrogeochimico dei Campi Flegrei (NA). 1st Convegno Nazionale dei Giovani Ricercatori in Geologia Applicata. 1, 293–409.
- Valentino, G. M., & Stanzione, D. (2004). Geochemical monitoring of the thermal waters of the Phlegraean Fields. Journal of Volcanology and Geothermal Research, 133(1-4), 261-289.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Responses to the reviewers

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Ass. Editor comments:

Dear author, the manuscript is interesting and the statistical approach is robust, nevertheless, there is a lack of coherence between the first part of the paper and the statistical approach that need to be solved. In my understanding, the Giggenbach equation is not appropriate for this data set and is not clear what is the true interest of its use in this manuscript. I suggest you a change of the title and focus on statistical approach to unravel the correlations among ions and to discuss ions sources under the complex hydrogeochemical scenario of the Campi flegrei area. Based on reviewer comments it is mandatory a deep revision of hydrogeological terms and hydrodynamics of the studied aquifer. I truly consider the manuscript interesting, but in the present version is not easy to read and to understand the exact logical of the document. The major surprise is the relationship of the Giggenbach equation with the robust statistical data analysis performed. Once again I suggest to make an effort and to separate these two parts omitting all the text related to equation.

I hope these comments will help you to reassess the manuscript.

The manuscript is revised based on the comments of the respected associate editor and reviewers to improve the quality. We hope that the revised version is acceptable for publication.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Reviewer #1:

This paper presents an analysis of methods used for grouping and classifying the compositions of fluids in geothermal areas, and shows how the use of hierarchical clustering approaches can offer a better way of identifying groups of fluids of similar origin in a complex setting, such as the Campi Flegrei, Italy. This is an interesting analysis; however the paper is written in a very dense way, which makes it very hard for the reader to follow. I have made a number of suggestions for improvement below - most of which relate to the presentation of information to the reader.

Thank you for your interest and for giving us an opportunity to improve quality of the manuscript.

In terms of the science, the paper lacks any discussion of sensitivity (e.g. of the clustering technique); and of uncertainties (how are they accommodated in this approach?).

In lines 325 to 330, it is explained that including or excluding only one variable can result in completely different hierarchical clustering results and the reason for considering the major anions for HCA is discussed. The HCA results are further supported with detailed hydrogeochemical studies in literature to ensure HCA performance.

Considering the 95% confidence interval in Fig. 4, it is also indicated that the difference in average Cl and HCO3 ratios between the mixed and bicarbonate–rich groundwaters is statistically significant (lines 502-504). The p-values are included in Fig. 8 as well to consider statistical significance of the linear regression models. Thank you for the constructive comment.

In terms of the 'worked analysis' can you explain to the reader how to go from the data to the analysis: what analysis package(s) do you use; where can the interested reader find the code? The cluster analysis and data transformation sections are very hard to follow.

A flowchart (Fig. 2) is added to clarify different steps and the packages (or the article shared the R code) are introduced in the manuscript (lines 311, 374, 392 and 402). The section regarding compositional data analysis is summarized as far as possible and the interested readers are referred to the article that explained details. The code for generating Fig. 10 is added to the supplementary material as well.

In terms of style, the use of abbreviations throughout the text makes the paper very hard to read. Most of the abbreviations are not needed!

Lines 424-429 are added at the beginning of the section 4 and we tried to use abbreviations less, as far as possible. We hope that the changes enhanced readability of the manuscript.

Title - the title of the paper has little or no meaning to the new reader. The 'modified .. diagram' isn't actually modified, isn't it just that you have modified the fields that plot within the diagram?

According to the comments from the respected associate editor, the Giggenbach diagram is not appropriate for the present dataset. Hence, the diagram is removed from the manuscript and the title along with the text is revised accordingly.

Highlights - should not contain acronyms or abbreviations that might not be widely understood (HCA? CoDA?)

The research highlights are revised.

Line 21-22 this sentence is not clear: what needs to inferred indirectly? What are the hybrid samples?

The abstract is rewritten because some figures are removed from the manuscript. We hope that everything is clear now.

Line 23 'the main information' - explain what you mean

The sentence is revised. Please see line 25.

Line 24 - delete the phrase (Euclidean . .. ..) - this should not be in the abstract.

Done.

Lines 24-30 : in an abstract you should summarise the main outcomes, the details of the methods can be explained in the paper

The abstract is upgraded.

Line 32 'and to save time' not needed in the abstract?

The phrase is removed from abstract (line 44).

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Line 33 'CoDA' please don't introduce new abbreviations in the abstract

The updated abstract does not contain abbreviations anymore.

Line 34 - 'this pioneering article' - not needed in an abstract! Let the readers decide.

The phrase is removed.

Line 56 'before 2007' What is the significance of 2007? Surely geothermal emanations at Campi Flegrei have been studied for centuries?

Agree. In this sentence (lines 76-81), it is highlighted that the studies on the entire study area were common before 2007, but researchers have focused on specific sectors of Campi Flegrei since 2007. The "Introduction" is rewritten according to the comments from the respected reviewers.

Line 126 'nevertheless' is not needed here

It is removed (line 171). Thank you.

Line 162 - water table variation: is this the depth below the surface?

The water table is 0 to 22 meters above sea level (Fig. 1) and a variation of 0.12 to 0.95 m is expected annually.

Line 198 - what assumptions do you need to make about the data in order to use HCA? How do you define 'homogeneous' 'similar' and so on? Do the spatial parameters influence the clustering, and how sensitive is the outcome to this parameter?

Further information is included in section 3.2.1, Fig. 4, Fig. S1, lines 467-469 and 502-504 to clarify these points. In this study, it is tried to obtain a representative dataset by collecting samples from different parts of the volcanic aquifer. The similarity (or dissimilarity) of the observations is measured by the distance matrix which is used for HCA. The dendrogram (Fig. 3a) is then cut based on the results of deep hydrogeochemical studies in the study area to identify the homogeneous groundwater groups. Geochemistry of the groups are finally investigated in the section 4.3 by the compositional data analysis technique to ensure that the groundwater samples are meaningfully categorized. In Fig. 8, for example, the p-values are added which show that the linear regressions are statistically significant. The authors appreciate this constructive comment.

Line 232 - please avoid using abbrevations in the text; it makes it very hard to read. (e.g. 'used CoDA and calculated ilr'

The abbreviations are removed (lines 361-364).

Line 240 - 'was compared' - is this what you have done in this paper? If you write these sentences in a more direct fashion, they will be easier to follow.

The sentence is revised (lines 371-374) for the sake of clarity.

Section 3.2.2.2 - the first paragraph (lines 244-256) is impenetrable, and completely unintelligible to the non-specialist reader.

In order to be more suitable for non-specialists, the section is revised and the interested readers are referred to the article that explained details. Please see lines 377-406.

Lines 299, 301 - Please explain *all* abbreviations when you first use them: BGW, CIGW. Better still, avoid using them when you can.

Majority of the abbreviations are introduced in lines 424-429 for further clarification.

Line 313 – how exactly does the HCA work? In Fig 3, what is the 'height' parameter? How sensitive is the analysis – if you omit samples randomly, do you end up with the same clustering? In Fig 3 can you label the well locations – there is no easy way to cross-compare with the dendrogram otherwise.

Further information about cluster analysis is provided in section 3.2.1 and the "height" parameter is defined in the caption of Fig. 3. Regarding the table provided below, excluding one sample (for example) leads to the same result, except in one case. Regarding the literature, HCA demonstrates the multivariate behavior of a dataset, but the results need to be further supported with detailed hydrogeochemical studies to ensure HCA performance. Because the previous studies in the study area support the results of HCA, the groundwater groups could be meaningful.

Good idea. The labels for sampling points are added to Fig. 3b to facilitate cross-comparison with the dendrogram.

Thank you for the constructive comments.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Comparison of HCA results applied on all samples vs. 43 samples (i.e. one sample is omitted randomly). In all examples, four clusters were considered.

Sample ID	HCA on all samples	One sample is randomly removed before HCA									
		Example 1	Example 2	Example 3	Example 4	Example 5	Example 6	Example 7	Example 8	Example 9	Example 10
W25 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W26 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W27 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W28 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	×	C1	C1
W29 (BGW)	C1	C1	C1	C1	C1	C1	×	C1	C2	C1	C1
W30 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W31 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W32 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W33 (BGW)	C1	C1	C1	C1	C1	×	C1	C1	C2	C1	C1
W34 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W35 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W36 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W37 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W38 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W39 (BGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W40 (BGW)	C1	C1	×	C1	C1	C1	C1	C1	C2	C1	C1
W41 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W42 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C2	C1	C1
W43 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W44 (BGW)	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
W3 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W4 (ClGW)	C2	C2	C2	×	C2	C2	C2	C2	C1	C2	C2
W5 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
W6 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W7 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W8 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W9 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W10 (ClGW)	C3	×	C3	C3	C3	C3	C3	C3	C3	C3	C3
W11 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
W12 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
W13 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W14 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
W15 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W16 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	×	C2
W17 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W18 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W19 (ClGW)	C3	C3	C3	C3	×	C3	C3	C3	C3	C3	C3
W20 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W21 (ClGW)	C2	C2	C2	C2	C2	C2	C2	C2	C1	C2	C2
W22 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	×
W23 (ClGW)	C3	C3	C3	C3	C3	C3	C3	×	C3	C3	C3
W24 (ClGW)	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3	C3
W1 (SGW)	C4	C4	C4	C4	C4	C4	C4	C4	C4	C4	C4
W2 (SGW)	C4	C4	C4	C4	C4	C4	C4	C4	C4	C4	C4
% of correctly clustered samples		100	100	100	100	100	100	100	56	100	100

Line 352 – please don't ask the reader to compare a figure from the supplementary data (S2) with one in the paper: it's impossible! Why not use them both in the main paper?

The figure is included in the manuscript. Please see Fig. 6.

Other typographical notes: 'McQuitty' not mcquitty throughout

Corrected. Thank you for the comment.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Reviewer #2:

Dear author,

Your manuscript presents a statistical approach to geochemical data acquired in the camps Flegrei fluids. The manuscript is relatively well written, but can be quite confusing in some part. I attached a hand reviewed version of your pdf, as well as the supplements.

I must admit I had to read more than once the manuscript just to get my head around what is the main message of the manuscript. I think the approach to "revise for modify the Giggenbach diagram" is misleading and detour the attention of the reader from the new and truly interesting statistical approach. You are representing the statistical composition ad HRA analyses in the Giggenbach diagram to link with previously acceptable analysis. Anyway based on that I would modify the title and the text to reflect more the statistical work.

The title and the text is modified based on the comments from the associate editor and the respected reviewers. Thank you for the constructive comments.

The introduction needs some work. It is a succession of separate paragraphs that do not have a cohesive link between them. Please rewrite.

The section "Introduction" is rewritten. We hope that the research gap and the objectives of this article are clear now.

Similarly the discussion needs some work; Fig 8 and 9 are barely discussed although they appear to be very interesting.

A couple of sentences are added to improve interpretation of the figures (lines 589-622). In addition, more information is included in Fig. 10 based on Fig. 1, Fig. 3 and Table S2. It is worth mentioning that these figures support the discussion in the previous sections and including more details lead to some repetition.

I think for the benefit of the reader and to open your reader and citation potential if you could add in plain language the benefit of each individual statistical approach. I am a little bit familiar with HCA, but I must admit I got really confused with the CoDA and all its ramification.

Based on the comments from the associate editor and the respected reviewers, the section 3.2.1 is improved. In addition, compositional data analysis is introduced in the last paragraph of introduction and discussed in lines 340-350 more. It is tried to express the advantage of this approach a bit further in sections 3.2.2.1 and 3.2.2.2, but the interested readers are referred to the articles developed the analysis for details. Thank you for the comment.

The figures need to be presented in a consecutive order and every panel needs to be labelled (Example Fig.6A). Avoid, top right..etc.

The panels are labeled in all figures to facilitate conveying message.

THE COMMENTS IN THE HAND-REVIEWED VERSION OF THE MANUSCRIPT:

I would modify the title to explain your novel approach than just explaining Giggenbach's diagram.

The title is modified as requested.

Line 22: Is 44 samples enough to have a good statistical approach?

Further information is provide in the section 4.2 for HCA. Calculating anion ratios alleviates the data problems for conducting hierarchical clustering. Regarding compositional data analysis, the log-ratio transformations are applied on each individual sample and no assumption about statistical distribution of data exists. Hence, it makes sense to apply HCA and compositional data analysis on 44 samples. In case the respected reviewer recommends using any particular test, we would be happy to improve the manuscript further.

Line 46: I think you can add recent Jolit et al. 2021.

Unfortunately, we could not find the article of Jolit et al. (2021), but we added another recent citation at the end of the sentence (please see line 60). If details of the above-mentioned article (e.g., DOI) is provided, we would include it in the article as well.

Line 46: Remove "dangerous".

Done (line 67).

Lines 71-75: Why this detail about As? Is there no other element directly affected by water-rock interactions? This appears a bit out of place.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Because concentration of As will be further discussed in the section "results and discussion", this information is included in the article from an investigation about mineral control of arsenic content in the thermal waters of Campi Flegrei. After rewriting the introduction, the paragraph is moved to the following section about the study area.

The relationship between water-rock interaction, CO2 fugacity, NaCl brine composition and pH buffering are also explained in lines 228-231, 451-454, 475-478, 493-494, 549-551 and 619-621.

Line 76: Remove "The interesting".

Done (see line 189).

Line 77: Map?

The sentence is placed in the following section (see line 189) and the corresponding figure is cited.

Lines 79-87: I think that most of this detailed chemistry information should be featured in the main text. In the introduction just lay out the key points to formulate your hypothesis. This can go in a subparagraph in study area.

As suggested, the "Introduction" is rewritten, some parts are summarized and some information is moved to the section "study area". Thank you for the constructive comment.

Lines 94-107: again, present your point in regard to your work. I think this can be rewritten to be more link to your problematic. By doing so, you will be able to better link your paragraphs.

We tried our best to rewrite the "Introduction". We think that the revised introduction provides some information about hydrothermal areas, Campi Flegrei, the importance of this research and the research gap that we would like to fill. Thank you again and please let us know if this section needs to be improved more.

Line 113: "long history"-But you worked mainly on your early dataset.

It is worth mentioning that hierarchical cluster analysis is an exploratory data analysis which is sensitive to excluding (or including) a single variable. Hence, the most important variables and the representative number of clusters are selected based on the long history of hydrogeochemical studies in the study area. These points are highlighted in Fig. 2 and section 3.2.1. Finally, HCA results are supported with detailed investigations in the past to ensure a reliable conclusion.

Lines 114-115: "hierarchical cluster analysis" and "compositional data analysis": This is what you need to emphasize. Upgrading the HCO3-SO4-Cl diagram is the extrapolation of the new statistical approach.

The introduction is generally rewritten and we hope that it is improved enough.

Lines 117-118: Remove the sentence.

Done.

Line 121: "polygenetic" – What do you mean? I would just delete.

It is deleted. Please see line 166.

Line 138: ". This"

Corrected (line 183).

Line 149: "metamorphic" – Do you mean hydrothermally altered?

Corrected (line 253).

Line 169: "saline waters" – Cl rang?

Chloride ranges from 6000 to 18000 mg/l.

Line 170: just 1?

Valentino and Stanzione (2003) considered average concentrations ions (determined in several samples) for identifying the end-members. The sentence is revised. Please see lines 240-241.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Line 172: so not a true group…

Four end-members were proposed and the remaining samples were considered as mixed waters. This part is revised. Please see lines 235-244. Thank you for the comment.

Lines 171-177: it would be good to give as well just the composition of the waters. Table?

The chemical data of Valentino and Stanzione (2003) and Aiuppa et al. (2006) are provided in Tables R1 and R2 at the end of this letter. The tables might distract attention of the readers from the main message of the article. However, if the respected reviewer recommends, they could be included in the supplementary material.

Line 195: Why not using as well published analysis? It would provide a greater point analysis for your statistical approach. Please explain your rational not using published data as well?

Some information is provided in lines 274-278. In addition, it is worth mentioning that chemical composition of geothermal waters and local seawater (reported by the previous studies) are used in most figures to understand evolution of groundwater. Thank you for the comment.

Line 197: Remove "for modifying the Giggenbach's diagram"

Done (line 296).

Line 198: Give example

Some examples are added (line 298).

Line 199: Repeat

The abbreviation "HCA" is introduced in lines 159 and 296. Hence, it is expected that the readers are familiar with this abbreviation in the middle of the manuscript.

Line 199: Remove "highly"

Done (line 299).

Lines 208-209: ?

The sentence is revised. Please see lines 328-330.

Line 211: more closer

Done (lines 331-333).

Line 215: Replace "Giggnbach's" with "ternary".

The Giggnbach's diagram is omitted from the manuscript as the respected associate editor requested.

Line 218: Fig. S1 – Maybe this needs to be moved in the main text.

The Giggnbach's diagram and the related charts are omitted from the manuscript as the respected associate editor recommended.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Line 227: I think because you are as well discussing Piper diagram, you have another reason for stepping out of the "modified Giggenbach's approach".

The Giggnbach's diagram and the related discussion are omitted from the manuscript as the respected associate editor recommended.

Line 238: more closer

Done (lines 365-369)

Line 239: Fig. 6 – out of order. Not sure how that relates.

Fig. 6 (i.e. Fig. 7 in the revised version) is the ilr-ion plot proposed by Shelton et al. (2018) as alternative of Piper diagram. By the way, it is removed from line 370.

Line 242: so is it your work or Shelton?

The sentence is revised. We hope it is clear that Shelton et al. (2018) proposed ilr-ion plot. Please see lines 371-375.

Lines 244-246: Please rewrite your leading sentence. I do not know what you are talking about and the link with the following …

The sentence is revised (lines 377-379). In addition, compositional data analysis and the advantages are explained in the last paragraph of "Introduction" and in lines 340-350.

Line 252: "data closure"?

More information is included (lines 388-389).

Line 253: why are you talking about marine sediment here?

The sentence is revised (lines 381-391).

Line 255-256: It would be great to add a "plain English language" line about the benefit of each statistical approach. This section is really hard to follow for less experts.

Thank you for the comment. The compositional data analysis and the benefits are discussed in the last paragraph of "Introduction" and in lines 340-350. Moreover, regarding the comments from reviewers #1 and #2, section 3.2.2.2 is rewritten (lines 377-406) and the interested readers are referred to the article that explained details.

Lines 244-256: This is a review of statistical approach and their problem. I suggest to move this out of your own statistical work to clarify your own study.

The paragraph is rewritten (lines 377-406), general information related to compositional data analysis are moved to the last paragraph of "Introduction" and to lines 340-350. The main idea of the paragraph is kept because the reader should know the reason for generating scatterplots with symmetric coordinates and log-transformed data, for example, in Fig. 8. However, the interested readers are referred to the publication that explained the details to improve readability.

Line 257: More closer

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

I think you need to explain more here.

Based on the comments from both reviewers, the equation is removed and the interested readers are referred to Kynčlová et al. (2017) for more information.

Lines 258-266: Honestly, this is beyond my skills. I hope the other reviewer can help. I have no idea what you are trying to do and for what purpose.

Based on the comments from reviewers #1 and #2, section 3.2.2.2 is rewritten and the interested readers are referred to Kynčlová et al. (2017) for more information.

Line 297: Why not plotting their analysis as well?

The Giggenbach's diagram is removed based on the comments from the respected associate editor. The data from Aiuppa et al. (2006) is not used for hierarchical clustering because enough information about each sampling point is not available and we could not verify if HCA groups the samples in a meaningfully. Please see lines 274-278.

Line 299: "BGW" – define

Line 301: "ClGW" – define

Line 303: "SGW" – define

Regarding these comments and the ones from reviewer #1, using abbreviation is avoided as far as possible. Some explanation is also provided in lines 424-429 for clarification. Thank you for the comment.

Line 313: "distance from the coast" – I am not sure what you mean. Fig. 3 does not reflect this.

The sentence is revised for the sake of clarity (lines 469-475).

Line 325: Fig. S1 – I would put this middle panel in the main text if you can.

Because the Giggenbach's diagram is omitted, Fig. S1 is irrelevant. However, bicarbonate, sulfate and chloride concentrations of the groundwaters collected in Puteolane well are included in Table S3.

Line 330: Your interpretation? Or reference missing

The reference is "Valentino and Stanzione, 2004" which is cited at the end of the next sentence (line 496) for avoiding repetition.

Line 333: replace "geothermal" with "hydrothermal".

Done (line 497).

Lines 335-336: Remove this sentence.

Done. Please see lines 500-501

Line 343: "other parameters" – List them.

Done (lines 518).

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Line 351: Include ClGW, BGW … in parenthesis.

Done (line 529).

Line 361: Please refer to the specific diagram.

Done (please see line 543).

Line 366: Fig?

Included (line 549).

Line 367: replace "introduced" with "proposed".

Done (line 515).

Lines 370-375: Reference to figure is needed.

The figure is cited (lines 552-558).

Line 376: A, B, C?

The corresponding figure part is cited (lines 559).

Line 387: A, B, C?

The citation is improved (line 574).

Line 394-395: You need more detail on that.

Thank you for the constructive comment. Fig. 7 (Fig. 8 in the revised manuscript) and the related discussion is upgraded. The p-values are added to the figure to represent significance of the results. Please see lines 575-578.

Lines 402-403: cite the figure at the end of the sentence.

Done (lines 596).

Line 409: remove "Moreover,".

Done (line 606).

Lines 412-413: Replace "lower-right" and "upper-left" with figure numbers.

Each quadrant of Fig. 10 are labeled and the sentences are modified as requested (lines 609-612).

Lines 429-433: Take the ample of new, more …… to chemistry. Because - - -

But you barely discussed this. Please put this in the main text.

The conclusion is rewritten because some figures are removed from the manuscript based on the comments from the associated editor and reviewers.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Line 451: "References"

Corrected (line 673).

Lines 621-664: Label the figure panels.

Done.

Line 653: Put the Piper diagram in the main paper for comparison.

It is included in the paper as Fig. 6.

THE COMMENTS IN THE HAND-REVIEWED VERSION OF THE SUPPLEMENTARY MATERIAL:

Label each panel in every figure and move Fig. S1b and S2 to the main text.

Fig. S1 is removed from the manuscript because the respected associate editor suggested omitting the Giggenbach's digram, but Fig. S2 is included in the manuscript as Fig. 6.

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 Analytical results of thermal waters from springs (s), wells (w) and pools (p) in Campi Flegrei (https://doi.org/10.1016/S0009-2541(02)00196-1). All concentrations are in mg/l, except for As, Hg, Tl and Pb (in μg/l).

Variable	CF 1(s)		CF 2(w)											
	1985	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994
T	55	55	57	57	57	57	58	58	44	39	45	31	46	45
pH	6.5	6.5	6.7	6.5	6.4	6.3	6.6	6.5	6.7	6.6	6.6	6.5	6.8	6.7
TDS	10,000	9000	6450	7725	8100	7770	8000	7500	3050	3016	3075	3300	3080	3075
Na	2060	1945	1960	2040	1970	2075	1660	1815	778	730	765	815	720	742
K	312	318	347	326	285	264	327	323	162	149	140	140	155	165
Ca	255	261	281	277	252	257	248	249	169	152	164	175	168	176
Mg	75	73	61	68	66	73	75	75	34	54	32	34	58	39
Fe	–	–	0.02	0.08	0.05	0.01	0.18	0.19	0.06	0.06	0.04	0	0.14	0.20
Al	–	–	–	0.92	–	0.10	0.82	0.84	–	0.62	–	0	0.54	0.52
Sr	–	–	1.8	2.3	1.9	0.4	0.1	0.8	0.7	0.9	0.4	0.3	–	–
Li	1.4	1.4	1.22	1.23	1.14	1.13	1.14	1.31	0.90	0.90	0.82	0.83	0.85	1.04
Rb	1.20	1.12	1.01	1.32	0.93	0.88	0.85	0.83	0.35	0.56	0.26	0.27	0.26	0.28
NH_4	–	–	6.7	2.6	5.7	4.7	8.7	8.4	4.1	2.6	3.9	2.9	0.3	0.7
Cl	3467	3335	2997	2900	2920	2910	2495	2790	465	310	350	450	485	520
HCO_3	1202	1220	1467	1612	1587	1586	1536	1525	1923	2020	1953	1979	1877	1842
SO_4	355	364	285	327	337	282	288	291	251	272	274	227	228	241
F	1.2	1.6	1.2	3.8	1.8	1.3	1.3	1.7	2.3	2.1	2.5	1.7	2.6	2.3
B	16.0	10.0	11.2	12.2	10.0	11.2	10.2	12.0	7.8	7.5	8.0	8.2	5.5	8.0
SiO_2	148	150	122	129	132	143	110	130	104	97	107	101	101	125
As	260	311	150	365	380	275	420	190	175	180	225	100	365	165
Hg	5	16	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	14.3
Tl	–	0.1	5.6	0.2	0.7	0.5	9.3	10	7.5	9.1	9.4	4.7	2.9	3.8
Pb	–	–	12.7	10.7	3.8	13.1	11.7	2.8	11.6	4.9	3.4	5.3	3.4	5.8

tr.: below detection limit (As=10 Ag/l, Hg=1 Ag/l, Tl=1 Ag/l, Pb=1 Ag/l).

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 (continued)

Variable	CF $3_{(W)}$								CF $4_{(W)}$						
	1985	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994
T	44	49	54	51	50	49	50	50	86	84	85	84	85	84	84
pH	6.8	6.9	6.7	6.7	6.7	6.6	6.7	6.8	7.1	7.3	7.0	7.3	7.0	7.3	7.2
TDS	2496	2300	2475	2962	2898	2738	2640	2493	3900	4350	4800	4950	4790	4900	4060
Na	667	699	810	720	700	725	685	705	966	1200	1295	1310	1300	1250	1195
K	127	135	184	149	142	130	145	167	260	487	390	356	165	407	423
Ca	24	28	52.1	52.1	36.1	24.1	36.1	40.1	32	44.1	44.1	40.1	44.1	36.1	48.1
Mg	9	19	9.7	9.7	9.7	9.7	9.7	9.7	5	2.4	4.9	6.4	4.9	4.9	4.9
Fe	–	–	0	0.02	0.05	0.02	0.26	0.38	–	0.14	0.07	0.28	0.17	0.19	0.54
Al	–	–	–	0.46	–	0.07	0.40	0.10	–	–	0.55	–	0	0.52	0.19
Sr	–	–	0.2	0.3	0.2	0.2	0	0	–	0.5	0.5	0.4	0.6	0	0
Li	0.47	0.21	0.61	0.58	0.49	0.48	0.48	0.63	0.7	0.89	0.82	0.73	0.84	0.74	0.87
Rb	0.28	0.244	0.32	0.53	0.14	0.14	0.14	0.16	1.6	1.4	1.41	0.99	1.20	1.12	1.14
NH_4	–	–	0.1	0	0.1	0.2	0.1	0.4	17.2	20.4	21	18.6	19	23.5	21.6
Cl	791.7	881.4	999	850	805	760	855	780	683	1240	1180	1300	1230	1380	1325
HCO_3	500	534	600	663	684	622	597	592	647	636	699	554	600	508	551
SO_4	280	230	218	217	164	150	167	218	1095	1020	1136	1148	1100	1067	1061
F	8.8	12.3	6.7	4.6	11.0	1.3	4.0	0.5	4.5	4.7	2.2	3.4	2.3	3.5	4.3
B	1.3	0.88	0.9	1.3	0.8	1.2	1.0	1.4	24	48.4	32.8	38.0	23.6	44.8	48.0
SiO_2	148	150	104	97	106	110	120	121	132	150	184	257	200	165	144
As	340	384	190	460	505	465	575	690	1630	3370	2255	2105	760	2735	1215
Hg	5	16	5.1	6.8	5.0	4.3	tr.	17.2	3	7.5	4.3	9.5	6.1	tr.	12.8
Tl	–	0.2	2.8	1.8	1	0.6	0.6	tr.	0.1	3.1	2.1	2.8	1.1	1.4	tr.
Pb	–	–	5.1	3.1	3.7	5.2	1.3	1.6	–	4	3.3	9.5	6.6	9.6	4.3

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 (continued)

Variable	CF 5(W)						CF 6(S)							
	1989	1/1994	3/1994	5/1994	9/1994	11/1994	1985	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994
T	70	36	55	70	72	68	88	76	74	75	74	75	76	75
pH	6.5	6.8	6.7	6.5	6.6	6.6	6.6	6.9	6.9	6.8	6.6	6.7	6.7	6.7
TDS	25,740	12,750	14,250	22,125	36,400	15,050	21,840	19,500	16,500	19,875	20,250	20,000	19,600	15,400
Na	7840	4870	4260	5860	9870	6500	7000	7000	6140	6500	5800	6000	7225	6645
K	188	201	181	220	718	272	330	240	192	194	166	170	197	189
Ca	293	245	204	280	545	260	311	271	224	240	232	240	236	240
Mg	46	114	92	95	666	122	185	146	34	58	39	40	36	29
Fe	–	0.13	0.80	0.08	0.41	0.27	–	–	0.06	–	0.07	0.08	0.33	0.23
Al	–	–	1.26	–	1.51	1.02	–	–	–	1.13	–	0.04	1.07	1.05
Sr	–	2.9	2.5	4.3	4	2.2	–	–	4.1	3.6	3.6	3.8	2.6	2.0
Li	7.33	5.31	6.01	8.91	4.07	8.05	–	10.5	9.2	9.3	8.5	8.7	6.5	9.6
Rb	2.2	1.04	1.32	1.93	1.86	1.36	–	2.46	1.89	2.21	1.72	1.95	1.72	1.55
NH_4	7.8	3.9	7.2	7.0	0.1	6.5	–	11	9.5	9.8	8.8	9.5	10.3	10.1
Cl	13,767	8000	6500	10,810	18,150	10,180	12,870	12,375	9650	10,100	10,200	10,050	11,560	10,400
HCO_3	295	478	457	376	890	410	410	260	281	299	300	305	281	273
SO_4	490	449	541	471	1598	422	629	495	359	377	375	369	300	322
F	5	5.7	0.8	2.4	4.5	7.5	7.8	7.3	7.6	4.0	3.6	6.8	8.0	7.8
B	26	13.4	14.4	21.6	20.0	15.6	61	27	20.4	22.8	24.8	24.0	20.0	22.0
SiO_2	138	76	88	121	121	126	109	110	106	123	118	119	109	132
As	3550	2310	3095	3755	3815	1235	5600	4740	5200	3760	5200	4500	6050	1610
Hg	8	tr.	tr.	tr.	tr.	tr.	1	10	tr.	tr.	tr.	tr.	tr.	tr.
Tl	2.8	5.5	10.7	16.2	23.3	tr.	–	7.84	18.4	19.9	11.6	12.2	19.2	8.0
Pb	–	8.3	6.2	8	6.9	13.6	–	–	15.9	6.4	6.9	10.4	11.9	17.9

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 (continued)

Variable	CF 7A(P)							CF 7B(P)					
	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994
T	90	85	85	85	87	90	87	91	84	90	90	95	92
pH	1.7	1.6	2.1	2.0	1.9	1.5	2.1	1.7	2.7	2.1	2.5	1.7	2.3
TDS	5460	5499	7650	6750	7350	7050	6300	6150	4050	5550	6000	9100	5880
Na	11	11.2	10	13.9	11.9	11.8	8.5	11.2	8.3	11.7	10.5	11.1	7.7
K	54	61.9	49.4	46.2	45.2	54.8	50.9	53.4	17.7	43.1	39.5	46.9	47.1
Ca	74	20.1	48.1	60.1	25.1	88.6	56.2	48.1	28.1	120	150	84.9	99.4
Mg	6	17.1	36.1	24.3	20.2	19.5	15.7	17.1	19.5	20.2	25.1	21.2	16.8
Fe	–	102	125	161	158	128.9	89.8	–	57.9	131	34	111	88.4
Al	–	–	65.94	55.6	38.5	53.1	52.6	–	23.61	49.7	19	42.1	44.2
Sr	–	0.06	0.06	0.04	0.05	0	0	0.08	0.05	0.4	0.2	0	0
Li	0.5	0.02	0.03	0	0.01	0.02	0.02	0.03	0.02	0	0.01	0.02	0.02
Rb	0.37	0.26	0.6	0.23	0.19	0.22	0.15	0.22	0.29	0.17	0.21	0.2	0.15
NH_4	545	680	620	787	700	858	623	680	508	790	690	1026	859
Cl	21	50	10	8	9	7	35	33	6	7	9	9	26
HCO_3	0	0	0	0	0	0	0	0	0	0	0	0	0
SO_4	3401	3430	3669	3121	3322	3300	3240	3092	1400	3121	2500	3399	3110
F	0.6	3.6	30.0	11.0	1.1	1.4	12.5	3.2	20.0	19.0	0.5	1.5	27.5
B	0.75	0.3	0.3	0.3	0.1	0.2	0.2	0.3	0.4	0.3	0.1	0.1	0.1
SiO_2	292	248	288	281	260	330	231	227	238	270	250	302	200
As	39	80	70	120	40	1880	240	70	45	85	60	415	295
Hg	150	185	125	115	175	45	114	115	40	90	160	60	232
Tl	2.94	6.2	6.1	4.4	6.2	3.7	2.5	8.1	4.6	2.5	5.8	2.7	2.9
Pb	–	29.1	14.0	8.2	8.2	5.7	9.3	20.0	8.7	8.5	14.1	6.8	13.9

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 (continued)

Variable	CF 8(w)								CF 45(w)			
	1985	1989	1/1994	3/1994	5/1994	7/1994	9/1994	11/1994	5/1994	7/1994	9/1994	11/1994
T	51	51	52	51	52	55	53	53	20	18	19	18
pH	7.2	7.6	7.1	7.0	7.5	7.2	7.1	6.9	7.3	6.8	6.9	7.2
TDS	7700	7722	8550	13125	10275	9600	10010	8000	936	936	936	940
Na	2225	1959	3170	3900	3190	3200	3170	2760	59.6	62.0	77.8	61.0
K	204	185	261	310	221	223	240	248	38.5	37.6	40.6	31.7
Ca	40	43	72	108	80	96	68	80	152	136	156	140
Mg	54	38	145	284	144	198	129	136	14.6	19.5	14.6	19.5
Fe	–	–	0.08	0.05	0.04	0.02	0.03	0.04	0.02	–	0.16	0.06
Al	–	–	–	0.68	–	0.24	0.06	0.08	tr.	tr.	tr.	tr.
Sr	–	–	0.5	1.0	0.4	0.9	0.3	–	0.39	0.31	0	0.32
Li	1.1	1.09	1.08	0.82	0.95	0.85	0.92	1.04	tr.	0.04	0.05	0.06
Rb	0.54	0.71	0.61	0.95	0.46	0.75	0.45	0.43	tr.	tr.	tr.	tr.
NH_4	1.9	0.03	0.6	0.5	0.9	0.8	0.2	1.1	0.41	0.09	0	0.05
Cl	2102	1650	3830	5200	3680	4000	3540	3450	74.0	80.4	69.4	67.0
HCO_3	2599	2520	2000	1919	1919	1950	2115	2141	322	328	330.6	316
SO_4	624	600	1000	1302	1038	1000	931	701	210	182	280	224.3
F	4.1	2.8	2.3	1.9	2.2	1.4	1.9	2.2	1.8	1.3	1.4	1.7
B	35	24	32.8	30.8	35.6	37.6	33.6	32.0	0.2	0.1	0.2	0.2
SiO_2	100	105	84	88	93	85	107	73	48	54	56	72
As	2600	2200	3090	1885	2015	2010	2225	855	27	20	43	25
Hg	5	16	tr.	tr.	tr.	tr.	tr.	tr.	1	0.7	–	4.2
Tl	–	0.23	7.5	10.2	5.0	7.8	10.3	8.1	0.1	1.5	tr.	tr.
Pb	–	–	9.8	4.2	9.9	8.3	4.6	10.1	6.9	4.7	4.9	5.5

NOTE: THE LINE NUMBERS, FIGURE NUMBERS AND TABLE NUMBERS IN RESPONSES REFER TO THE MANUSCRIPT WITH CHANGES MARKED.

Table R1 (continued)

Variable	CF 50(w)				CF 20(w)			CF 15(w)			
	5/1994	7/1994	9/1994	11/1994	5/1994	9/1994	11/1994	5/1994	7/1994	9/1994	11/1994
T	–	–	–	–	15	15	13	18	17	17	16
pH	8.0	7.2	7.1	7.9	7.6	7.4	8.1	6.9	7.1	7.2	7.7
TDS	585	663	624	700	546	562	585	507	499	499	520
Na	100	110	120	109	110	102	116	70.3	80.2	78.4	72.4
K	26.7	28.6	30.4	24.2	22.4	26.7	16.1	18.5	19.6	21.7	16.7
Ca	32.1	36.1	32.1	32.1	24.0	36.1	12.0	28.0	20.0	40.1	36.1
Mg	4.8	8.2	9.7	9.7	12.2	4.8	4.9	9.7	12.2	12.2	12.2
Fe	tr.	–	0.12	0.08	0.02	0.17	0	0.01	–	0.15	0.28
Al	tr.	0.1	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.
Sr	0.09	0.11	tr.	tr.	0.1	tr.	tr.	0.09	0.14	0.09	tr.
Li	tr.	0.04	0.05	0.06	tr.	0.08	0.13	tr.	0.05	0.05	0.06
Rb	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.	tr.
NH_4	0.32	0.13	tr.	0.04	0.27	0	0.05	0.04	tr.	tr.	tr.
Cl	74.0	67.6	62.8	72.4	66.6	69.8	70.3	50.0	56.8	51.6	58.0
HCO_3	212	300	303	304	250	200	236	216	238	239	208
SO_4	71.9	49.3	66.8	41.0	79.9	83.4	51.2	62.2	45.0	57.0	43.2
F	1.4	4.2	3.8	5.2	2.8	0	1.7	0.8	2.6	2.5	3.2
B	0.3	0.3	tr.	0.4	0.4	tr.	0.3	0.3	0.1	tr.	0.3
SiO_2	43.3	43.2	46.7	52.2	39.8	44.8	38.9	39.8	52.2	45.1	57.8
As	16.8	21.0	31.1	27.1	45.0	–	35.8	24.5	18.3	30.6	35.3
Hg	tr.	tr.	tr.	11.4	1.3	–	tr.	1	0.9	tr.	15.8
Tl	2.4	1.6	0.3	0	1.3	–	tr.	tr.	2.6	0.9	6.3
Pb	5.6	3.9	2.4	4.3	7.1	tr.	3.3	12.8	4.5	2.8	1.9

Table R2 Chemical composition of Campi Flegrei groundwaters (https://doi.org/10.1016/j.chemgeo.2005.11.004)

	T	pH	Eh	Na	K	Ca	Mg	Cl	SO4	HCO3	NH4	B	Li	As	δ18O
ClGW															
1	85.3	7.3	0.116	1448.0	442.0	48.0	n.m.	1687.0	1076.0	589.0	21900.0	n.m.	1070.0	3610.0	n.m.
2	26.6	7.1	0.197	217.0	65.0	126.0	34.0	318.0	257.0	244.0	22.0	100.0	47.0	8.0	− 5.7
3	73.0	6.6	0.047	5961.0	170.0	284.0	49.0	8882.0	412.0	329.0	9511.0	24248.0	9783.0	6939.0	− 2.2
4	26.8	7.6	0.370	586.0	236.0	48.0	8.0	553.0	449.0	531.0	38.0		548.0	375.0	− 3.3
5	14.8	6.9	0.280	144.0	62.0	155.0	46.0	267.0	252.0	304.0	113.0	596.0	12.0	6.0	− 5.9
6	25.7	7.4	0.296	992.0	60.0	60.0	60.0	1520.0	84.0	421.0	25.0	3307.0	1200.0	378.0	− 6.7
7	47.8	6.8	0.115	592.0	112.0	23.0	4.0	579.0	174.0	580.0	100.0	1237.0	438.0	897.0	− 6.2
8	18.9	7.2	0.156	340.0	62.0	82.0	22.0	388.0	124.0	409.0	61.0	693.0	256.0	32.0	− 5.9
9	48.7	7.2	0.092	2038.0	175.0	54.0	62.0	2194.0	598.0	1919.0	615.0	25316.0	1005.0	2465.0	− 5.4
10	57.4	6.4	–	2022.0	296.0	259.0	73.0	2923.0	369.0	1473.0	12600.0	12836.0	1341.0	459.0	− 4.7
SGW															
11	94.0	1.7	0.046	12.0	53.0	93.0	13.0	14.0	3819.0	n.d.	613173.0	477.0	19.0	41.0	3.4
12	46.8	1.6	0.095	9.0	26.0	14.0	2.0	14.0	2375.0	n.d.	128856.0	362.8	10.0	128.0	8.1
13	95.1	2.6	–	25.0	34.0	315.0	8.0	5.0	1758.0	98.0	214263.0	22040.0	22.0	785.0	1.0
BGW															
14	37.6	6.3	–	798.0	150.0	119.0	36.0	439.0	297.0	1907.0	6421.0	7452.0	986.0	784.0	− 5.5
15	27.5	6.2	–	731.0	139.0	154.0	39.0	604.0	244.0	1668.0	1211.0	5896.0	779.0	36.0	− 5.1
16	30.9	6.3	0.080	655.0	135.0	118.0	40.0	409.0	287.0	1525.0	1313.0	7944.0	823.0	105.0	− 5.2
17	23.7	6.0	0.026	451.0	101.0	119.0	30.0	316.0	216.0	1130.0	2953.0	5692.0	484.0	119.0	− 5.7
18	24.6	6.5	–	244.0	55.0	96.0	21.0	98.0	89.0	870.0	361.0	802.0	244.0	72.0	
19	20.9	5.9	–	362.0	86.0	83.0	30.0	236.0	240.0	787.0	3655.0	7844.0	549.0	126.0	− 5.8
20	17.1	7.0	− 0.05	48.0	27.0	70.0	9.0	34.0	125.0	159.0	54.0	55.0	59.0	14.0	n.m.
21	24.7	8.1	0.130	249.0	7.0	17.0	1.0	143.0	160.0	222.0	40.0	1512.0	81.0	107.0	− 5.9
22	24.5	7.2	0.319	220.0	29.0	40.0	9.0	168.0	189.0	275.0	69.0	994.0	199.0	106.0	− 6.0
23	17.9	6.8	0.254	269.0	62.0	105.0	46.0	279.0	286.0	448.0	n.m.	1203.0	37.0	10.0	n.m.
24	26.9	6.6	0.050	241.0	43.0	58.0	2.0	176.0	228.0	415.0	58.0	181.0	63.0	27.0	n.m.
25	15.4	6.8	0.302	108.0	42.0	170.0	40.0	186.0	241.0	455.0	45.0	195.0	6.0	6.0	− 5.7

Units: T (°C); Eh (V); major ions (mg/l); NH4, B, Li and As (µg/l); δ18O (‰) versus V-SMOW.
n.m.: not measured; n.d.: below detection limit

Table R2 (continued)

	T	pH	Eh	Na	K	Ca	Mg	Cl	SO4	HCO3	NH4	B	Li	As	δ18O
26	33.9	8.2	0.213	426.0	89.0	11.0	n.d.	226.0	276.0	558.0	n.m.	487.0	314.0	772.0	− 6.1
27	20.4	7.5	0.108	58.0	33.0	78.0	7.0	41.0	129.0	248.0	65.0	95.0	55.0	43.0	− 6.1
28	17.5	8.2	0.134	260.0	11.0	30.0	3.0	166.0	121.0	284.0	113.0	174.0	243.0	25.0	n.m.
29	41.0	6.6	0.150	517.0	61.0	21.0	3.0	343.0	232.0	570.0	44.0	962.0	432.0	219.0	− 6.0
30	24.8	8.6	0.134	145.0	13.0	10.0	0.0	53.0	90.0	210.0	33.0	162.0	60.0	78.0	− 5.9
31	16.8	7.2	0.081	67.0	38.0	72.0	5.0	74.0	82.0	213.0	175.0	48.0	54.0	6.0	n.m.
32	22.6	7.4	0.232	211.0	14.0	44.0	12.0	177.0	120.0	355.0	99.0	536.0	157.0	50.0	− 5.7
33	20.7	7.4	0.316	235.0	37.0	74.0	20.0	205.0	131.0	415.0	7.0	547.0	98.0	49.0	n.m.
34	18.0	7.9	0.300	158.0	16.0	14.0	1.0	69.0	63.0	198.0	77.0	496.0	122.0	64.0	n.m.
35	17.6	7.0	0.194	194.0	68.0	160.0	41.0	167.0	215.0	641.0	28.0	850.0		11.0	− 6.2
36	21.5	7.2	0.256	71.0	41.0	133.0	20.0	66.0	135.0	403.0	32.0	109.0	87.0	67.0	n.m.
37	16.3	7.4	0.380	71.0	47.0	66.0	8.0	51.0	90.0	276.0	41.0		97.0	41.0	− 6.2
38	16.0	7.6	0.060	83.0	22.0	40.0	6.0	44.0	65.0	204.0	57.0	66.0	85.0	36.0	− 6.1
39	22.2	7.6	0.180	173.0	24.0	26.0	4.0	93.0	73.0	293.0	45.0	341.0	114.0	70.0	− 5.9
40	24.0	7.2	0.285	85.0	46.0	57.0	8.0	80.0	66.0	268.0	86.0	213.0	33.0	26.0	n.m.
41	25.1	8.7	0.137	178.0	12.0	11.0	n.d.	74.0	60.0	271.0	43.0	730.0	109.0	60.0	n.m.
42	17.8	7.2	0.247	79.0	50.0	74.0	6.0	59.0	79.0	324.0	53.0	112.0	85.0	48.0	− 6.2
43	20.8	6.7	0.256	76.0	42.0	76.0	15.0	74.0	60.0	278.0	54.0	404.0	29.0	12.0	n.m.
44	19.9	7.1	0.320	109.0	40.0	46.0	7.0	78.0	66.0	317.0	84.0	317.0	57.0	20.0	− 6.2
45	16.5	6.8	0.239	97.0	59.0	129.0	26.0	118.0	101.0	493.0	105.0	84.0	1.0	6.0	n.m.
46	21.9	7.8	0.105	216.0	32.0	31.0	6.0	97.0	85.0	421.0	51.0	789.0	87.0	77.0	− 6.1
47	20.4	8.4	− 0.246	145.0	15.0	15.0	n.d.	69.0	41.0	236.0	173.0	402.0	100.0	88.0	− 6.1
48	18.1	6.9	0.017	90.0	54.0	91.0	21.0	93.0	80.0	433.0	227.0	400.0	36.0	7.0	− 6.5
49	19.8	7.9	0.333	143.0	21.0	21.0	1.0	77.0	48.0	292.0	91.0	n.m.	91.0	71.0	− 6.0
50	18.5	7.0	0.252	69.0	46.0	81.0	17.0	76.0	67.0	366.0	1018.0	262.0	34.0	9.0	n.m.

Table R2 (continued)

	T	pH	Eh	Na	K	Ca	Mg	Cl	SO4	HCO3	NH4	B	Li	As	δ18O
51	15.2	6.8	0.295	61.0	38.0	92.0	19.0	80.0	59.0	356.0	265.0	90.0	4.0	8.0	− 6.1
52	26.5	8.6	0.074	145.0	13.0	12.0	n.d.	62.0	32.0	222.0	63.0	431.0	106.0	95.0	− 6.1
53	18.3	7.0	0.429	51.0	27.0	60.0	4.0	25.0	45.0	250.0	36.0	48.0	54.0	38.0	n.m.
54	18.7	6.7	0.130	100.0	51.0	88.0	24.0	124.0	49.0	441.0	230.0	504.0	10.0	13.0	− 6.6
55	18.7	8.3	0.154	57.0	9.0	40.0	2.0	51.0	15.0	161.0	141.0	32.0	7.0	2.0	− 5.6
56	17.2	7.2	0.262	217.0	65.0	126.0	34.0	26.0	28.0	198.0	24.0	85.0	31.0	29.0	− 6.2
57	22.6	8.0	− 0.073	210.0	27.0	15.0	3.0	131.0	5.0	369.0	38.0	816.0	61.0	23.0	− 6.1
58	18.9	8.2	0.250	155.0	3.0	10.0	n.d.	40.0	24.0	275.0	60.0	187.0	52.0	36.0	n.m.
59	14.7	7.2	0.200	125.0	26.0	27.0	9.0	37.0	33.0	357.0	59.0	651.0	35.0	44.0	− 7.4
60	20.3	6.7	0.285	113.0	117.0	79.0	30.0	65.0	37.0	664.0	1.0	83.0	16.0	25.0	− 7.3
61	21.9	6.6	0.211	214.0	32.0	29.0	5.0	157.0	254.0	206.0	63.0	226.0	10.0	11.0	− 5.6
62	38.2	7.6	0.190	363.0	85.0	47.0	7.0	290.0	223.0	393.0	14.0	369.0	395.0	404.0	n.m.
63	22.9	7.3	0.208	356.0	84.0	82.0	13.0	359.0	159.0	537.0	66.0	917.0	372.0	272.0	− 6.4
64	30.4	8.1	0.169	360.0	83.0	37.0	8.0	257.0	206.0	387.0	n.m.	362.0	345.0	404.0	− 6.2