University of Naples Federico II

Department of Earth, Environmental and Resources Sciences (DiSTAR)

Ph.D. in

Earth, Environmental and Resources Sciences
XXXIV CYCLE

Ph.D. Thesis

Geochemical study of different environmental media for characterization of the hydrothermal system and the subaerial environment of Phlegrean Fields, and the surrounding areas

Candidate:

Pooria Ebrahimi

Tutor:	Co-tutors:
Prof. Stefano Albanese (UNINA)	Prof. Vincenzo Allocca (UNINA) Dr. Stefano Caliro (INGV)

2018-2021

ABSTRACT

The Phlegrean Fields are located in southern Italy and the latest volcanic eruption in this area goes back to 1538 CE. In the last decades, the bradysismic crises were sometimes followed by evacuation orders from the authorities to manage an emergency. On the other hand, the study area is highly populated and urbanization, industries together with agricultural practices may pose a threat to the ecosystem and human health. The main objective of this thesis is to use some recently developed data treatment approaches (i.e. compositional data analysis and spatial analysis) for highlighting the dominant processes controlling geochemical composition of different environmental compartments.

The groundwater samples could be categorized in four groups: (1) the bicarbonate-rich groundwater mainly derived from poor-arsenic meteoric water; (2) the chlorine-rich groundwater with significant thermal/seawater contribution; (3) the sulfate-rich groundwater generated by the interaction of bicarbonate-rich groundwater and hydrothermal vapors around Solfatara volcano; and (4) the mixed groundwater observed where the three groundwater groups undergo a mixing process. Based on another prospecting campaign for determination of the dissolved radon and carbon dioxide (in addition to the major ions, lithium, arsenic, boron and physicochemical parameters), two geochemical associations are identified using multivariate statistics: (1) $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{K^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{HCO_3^{-}}$+$\mathrm{CO_2}$ and pH; and (2) $\mathrm{Na^{+}}$, $\mathrm{Cl^{-}}$, As, B, Li, Rn, TDS and T. It highlights that the groundwater composition is generally influenced by: (1) meteoric water, which is modified by the CO_2–rich magmatic gas in some cases; and (2) hydrothermal fluid and/or seawater.

Regarding Naples tap water, seasonal variation is observed in chemistry of the water exploited from the nearby aquifers and well fields. It stresses the importance of interconnections between urban water reservoirs for adjusting water quality, being complicated in cities like Naples with an old distribution network. Thus, the compositional nature of data is considered in hierarchical clustering to study water transfer between urban water reservoirs. The proposed method potentially defined the source of tap water in each city area based on chemical data and provided an overview of the relationships between different components of the network for encouraging advanced simulations.

The eastern sector of Pozzuoli Bay is divided into the proximal and distal zones for conducting research on the spatial distribution and source of elements in the marine sediment. In front of the former industrial site (i.e. the proximal zone), Hg, Cd, Cu, Pb and Zn are positively correlated with mud and organic matter. Concentrations of these metals decrease along the cores and in the distal zone. Nevertheless, Al, As, V, Fe, Cr, Ni and sand form an association along the coast which strengthens with increasing distance from fumaroles in the proximal zone. It suggests that the pyroclastic deposits of Phlegrean Fields and some of the seepages with a hydrothermal component are related to this group of elements.

Soil geochemistry indicates that higher Pb, Zn, Cd, Cr, Hg, Ni and Sb levels correspond to the urban area, but greater As, Tl, Co, Cu, Se and V quantities occur in other parts of the study area. The Pb–Zn–Hg–Cd–Sb–Cr–Ni association probably highlights anthropogenic activities in the urbanized area while the Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association mostly reveals the contribution of pyroclastic deposits. The Na–K–B association feasibly indicates the weathering degree, being lower in the Neapolitan Yellow Tuff caldera. The probabilistic health risk assessment for the children under 6 years old shows negligible non-carcinogenic risk for exposure to Pb and Zn through soil ingestion (certainty > 95%). Unexpected Pb carcinogenic risk through the ingestion route is detected for all age groups (CR ≤ 1E-06 with above 99% certainty). However, for the inhalation pathway, there is a 90% chance of acceptable (i.e. between 1E-6 and 1E-4) Pb carcinogenic health risk for the children aged <1 year old and a 25-45% chance for those between 1 and 6 years old.

Overall, the underground water is mainly influenced by the hyrothermal system although the sediment and soil geochemistry is affected by both natural and anthropogenic activities. The effectiveness of compositional data analysis for investigating groundwater hydrogeochemistry in the volcanic aquifer of Phlegrean Fields indicates that this technique deserves further attention for developing new tools and identifying promising precursors of volcanic eruption worldwide. The number of magnetic resonance imaging centers administering contrast agents for diagnostic imaging in the study area highlights that gadolinium contamination in water resources is a question to be answered in Phlegrean Field and the immediate surroundings. This line of research might also be advantageous in hydrogeological studies. Considering the environmental aspects, the results in this thesis should not be overlooked because Naples is a heavily populated metropolis under high environmental pressure and previous studies reported the increased Pb and Zn in soil over a 26-year timespan. In addition, the Bagnoli industrial site is still a source of pollution. Implementing a combination of spatial analysis, multifractal analysis, compositional data analysis and probabilistic health risk modeling helped to characterize the environmental status in detail.

PREFACE

This thesis is a “thesis with publication” which consists of seven chapters (including introduction and conclusions). It is based on five publications and three manuscripts (one of them is in preparation while the rest are under evaluation for publication). Some results were also presented in the Goldschmidt and BE GEOSCIENTISTS conferences in 2021.

To better organize structure of the thesis, title of each chapter mainly refers to a specific environmental compartment. More than one publication/manuscript is accordingly placed in some chapters. Specifically, chapter two is composed of two sections about the groundwater in Phlegrean Fields aquifer. The next three chapters are made of only one publication/manuscript each, but chapter six presents three scientific collaborations in unique sections.

It is worth mentioning that the supplementary materials of the publications are included in the corresponding chapter, the figure/table numbers are upgraded and some minor grammatical modifications are also applied. The content of chapter six is summarized to keep the main focus of the present document on the aims and objectives of the thesis. Finally, a unique references list is placed on page 201.

TO MY FAMILY

FOR UNCONDITIONAL SUPPORT

ACKNOWLEDGMENTS

I would like to begin by acknowledging Italians and pay my respects to them because of their kindness and hospitality. Regardless of the COVID-19 pandemic, I am happy that I have had the chance to visit Italy and live with friendly people. Trying pizza, pasta and coffee is enjoyable here and you could be proud of yourself. Living inside and around volcanoes is threatful, but it provides unique opportunities for geologists (including myself) to deepen their understanding of the earth and save lives in an emergency.

I would like to express my sincere gratitude to my Ph.D. tutor, Prof. Stefano Albanese, for the continuous support, critical comment and encouragement that significantly helped learning new approaches for geochemical data analysis, getting familiar with the geochemical processes in the study area and bringing this project to fruition. Thanks are also extended to the co-tutors, Prof. Vincenzo Allocca (UNINA) and Dr. Stefano Caliro (INGV, Napoli), for their valuable contribution and sharing their experience of the complex hydrogeological system in Phlegrean Fields. Overall, the tutor and co-tutors demonstrated what a brilliant and hard-working scientist can accomplish.

Furthermore, I appreciate conducting research with Prof. Domenico Cicchella (University of Sannio), Prof. Benedetto De Vivo (retired from UNINA; current affiliation: Pegaso Online University) and Prof. Annamaria Lima (UNINA) because I learned a lot from them during my Ph.D. program. Collaboration with a range of researchers at different universities/institutions such as INGV Napoli, CNR-IRISS (Napoli), CNR-INO (Pozzuoli), ABC Napoli, University of Florence, Sapienza University of Rome and Guilin University of Technology (China) played a significant role in this Ph.D. program as well. Moreover, taking part in the Erasmus+ learning agreement and working with the researchers at the Center for Ecological-Noosphere Studies (Armenia) improved quality and scientific impact of this thesis.

I would like to sincerely thank Annalise Guarino, Antonio Aruta, Andrea Gramazio and Lucia Pacifico (DiSTAR, UNINA) along with Salvatore Dominech (Tongji University, China) for their invaluable assistance during my study. The insightful comments from Dr. Daniela Zuzolo (University of Sannio) on a manuscript also helped to upgrade the work and speed up the publication process. Due to the outbreak of COVID-19, the achievements in this thesis were impossible without significant group work and I owe my deepest gratitude to everybody who provided the motivation of thinking critically and being productive in every step. Please accept my apologies that I could not mention every individual one by one. Finally, I wish to pay my special regards to my parents for going through lots of challenges to ensure my happiness and success. My genuine thanks to my siblings for being always available as friends. I owe you this milestone in the path of life.

TABLE OF CONTENTS

CHAPTER 1

Introduction

Introduction

The “Phlegrean Fields”, “Phlegraean Fields” and “Campi Flegrei” are interchangeably used in the literature and refer to a large active volcanic complex in south Italy (Fig. 1a) with a mainly explosive (from phreatomagmatic to Plinian) eruptive history of at least 80 ka (Scarpati et al., 2013; Peccerillo, 2020). The word “Phlegrean” derives from the Greek flègo “burn” which explains the importance of fumaroles and thermal waters since ancient times. The partially submerged caldera is not cone-shaped and the bradysismic crisis is the corresponding characteristic feature.

The (inter)national scientific community has been intensively investigated structural setting of the caldera, history of the volcanic activities, representative precursors of the volcanic events and the risk of volcanic eruption through tephrostratigraphy, tephrochronology, geochemical and geophysical monitoring along with drillings in the onshore/offshore sector of Phlegrean Fields. The majority of researchers proposed the 40-ka Campanian Ignimbrite (CI) and 15-ka Neapolitan Yellow Tuff (NYT) eruptions together with the associated collapses as the main factors being responsible for the present structure of caldera (e.g., Rosi and Sbrana, 1987; Fisher et al., 1993; Orsi et al., 1996; Di Vito et al., 1999; Perrotta et al., 2006; Natale et al., 2021). In contrast, Scandone et al. (1991), De Vivo et al. (2001), Rolandi et al. (2003) and De Natale et al. (2014, 2016) considered only NYT and Sbrana et al., (2021) proposed only CI as the caldera-forming eruption in Phlegrean Fields. Determining the caldera rims has also been a matter of debate over the last decades (Fig. 1b to e). Similarly, the hydrological and hydrogeochemical studies suffer from lack of a clear insight into structure of the volcanic aquifer (AGIP, 1987; Rosi and Sbrana, 1987; Dall'Aglio et al., 1972; Baldi et al., 1975; Cortecci et al., 1978; Bolognesi et al., 1986; Ghiara et al., 1988; Caprarelli et al., 1997).

The environmental issues also deserve further investigation in the study area. More than 2 million people live in Phlegrean Fields caldera and the immediate surroundings. The intensive industrial activity in the twentieth century led to adverse environmental impacts and raised concerns about ecological and human health risks (Bergamin et al., 2005; Morroni et al., 2020). The industrial plants were dismantled at the end of the last century, but some researchers still consider the coastal landfill as a source of pollution (Romano et al., 2009) and the authorities are organizing remediation plans to restore the brownfield and the coastal environment. Urbanization, agricultural practices and volcanic activity could be the other sources of potentially toxic elements in the study area (Grezzi et al., 2011; Bove et al., 2011). Several researchers (e.g., De Vivo et al., 2016; Petrik et al., 2018a; Thiombane et al., 2019; Guagliardi et al., 2020; Zuzolo et al., 2020) studied soil geochemistry for characterizing contribution of the anthropogenic and geogenic factors at a regional scale.

Overall, planning more drilling projects and geophysical surveys might help update understanding of the scientific community, answer the ongoing questions and validate the solutions. On a daily to an annual basis, a wide range of information is available today and size of the archived datasets is growing noticeably. The recent advances in data mining could also be advantageous to address some of the scientific challenges.

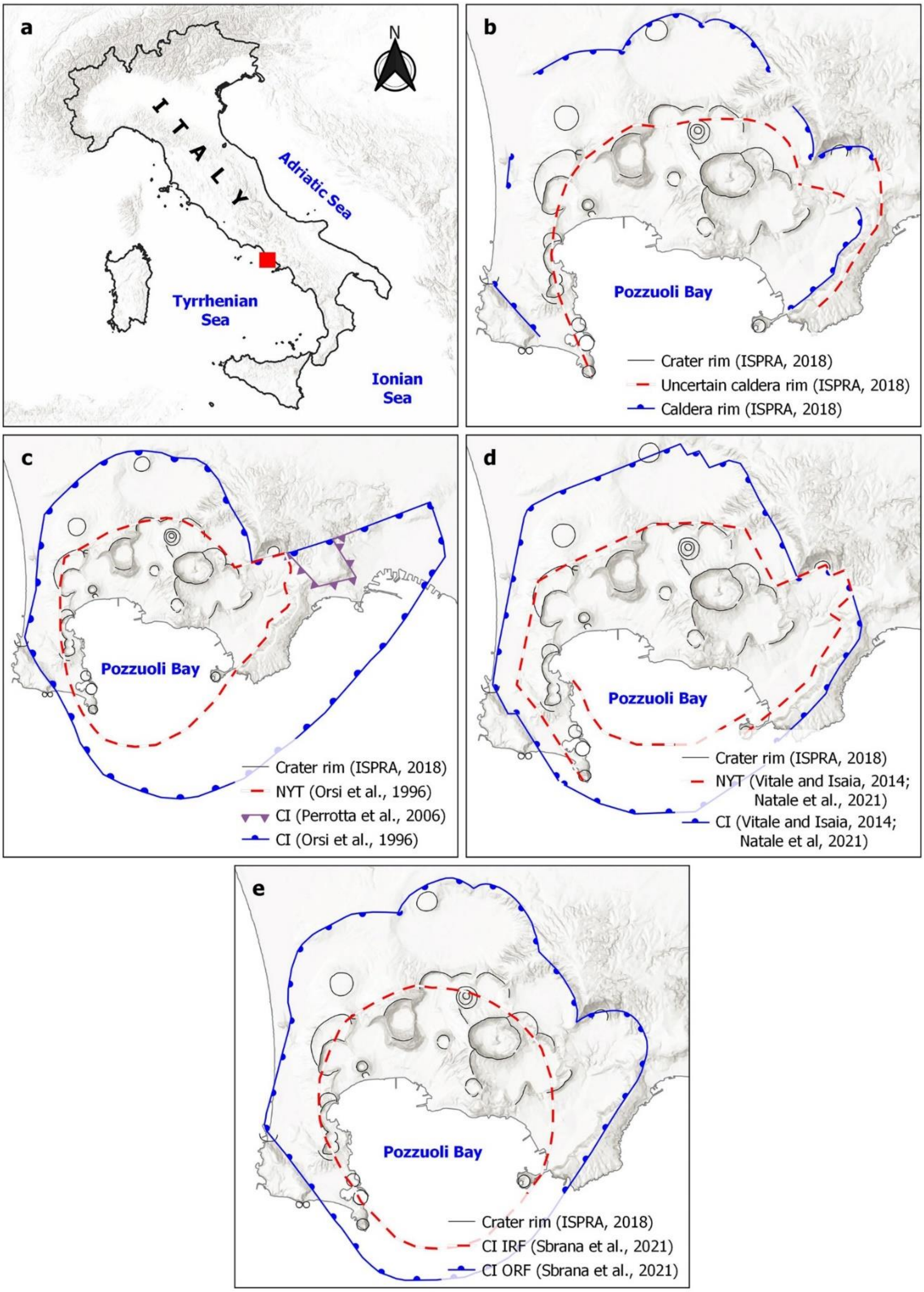


Fig. 1 (a) Location of Phlegrean Fields and (b to e) the proposed caldera rims by different researchers. NYT and CI refer to Neapolitan Yellow Tuff and Campanian Ignimbrite, respectively. IRF: inner ring fault system; ORF: outer ring fault system.

1. Aims and objectives

The main objective of the thesis is to use some recently developed data treatment approaches for highlighting the dominant processes controlling geochemical composition of different environmental compartments in the study area. Briefly, the following points are addressed in the thesis: (1) groundwater hydrogeochemistry in the Phlegrean Fields volcanic aquifer; (2) the contribution of natural and anthropogenic factors in composition of the sediments from Pozzuoli Bay – the submerged sector of Phlegrean Fields; and (3) topsoil geochemistry in Phlegrean Fields and the immediate surroundings. As a side activity focusing on the statistical techniques in the thesis, the tap water chemistry of Naples was also investigated during the first phase of mobility restrictions due to the COVID-19 pandemic.

2. Methodology

The methods applied for investigating each environmental media are comprehensively discussed in the corresponding chapter. However, a summary is provided here to give an overview to readers.

2.1. Datasets, sampling and chemical analysis

The (geo)chemical datasets for tap water, marine sediment and soil are provided by the responsible staff from ABC Napoli, the researchers from Vesuvius Observatory (INGV, Osservatorio Vesuviano, Napoli) and the tutor, respectively. The groundwater data are, however, produced in the Ph.D. program and will be briefly explained here.

In May 2019, a sampling campaign was conducted and groundwater samples were collected (Fig. 2) from various water types in the study area to capture influence of volcanic activity on groundwater composition and understand Phlegrean Fields hydrothermal system better. The water samples were filtered with 0.45 μm membrane filters into high-density polyethylene (HDPE) bottles for instrumental analyses. The aliquot for determination of the major cations was acidified (1%) with Suprapure 36% HCl, but the one for minor and trace elements was acidified (1%) with Suprapure 63% HNO_3. Samples were analyzed for Li, $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{Na^{+}}$, $\mathrm{K^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{Cl^{-}}$, $\mathrm{NO_3^{-}}$ and $\mathrm{F^{-}}$ concentrations by ion chromatography standard technique at the INGV-Osservatorio Vesuviano, Napoli. Arsenic and boron were detected using inductively coupled plasma-optical emission spectrometry (ICP-OES) at University of Naples Federico II (Fig. 3). Alkalinity, electric conductivity (EC), pH and temperature were determined with portable instruments in situ (Fig. 2).

From November to December 2020, another set of representative groundwater samples was collected (Fig. 2) in the study area for radon measurement via AlphaGUARD (viz the AquaKIT accessory) and carbon dioxide determination with ECOPROBE 5 at University of Naples Federico II (Fig. 4). At each station, two unfiltered and unacidified aliquots were collected in airtight sealed containers for measurement of the dissolved gases. Concentrations of ions and elements were also detected as explained in the previous paragraph, except for laboratory measurement of alkalinity in an unfiltered and unacidified aliquot on the same day of sampling. Instead of EC, total dissolved solids (TDS) were determined in this campaign with a portable device (Fig. 2). All containers were made of high-density polyethylene, but the aliquot for Rn was stored in a glass bottle. It was a big challenge to set appointments with owners of the water wells during the period of COVID-19 pandemic. However, we tried to provide the latest information on the volcanic gases in the groundwater body because some detailed research projects will start soon in the study area and identifying the targets plays an important role in filling the existing knowledge gap.

Fig. 2 Groundwater sampling (a to c) and field measurements: (d) alkalinity; (e) pH; (f) TDS; and (g) temperature.

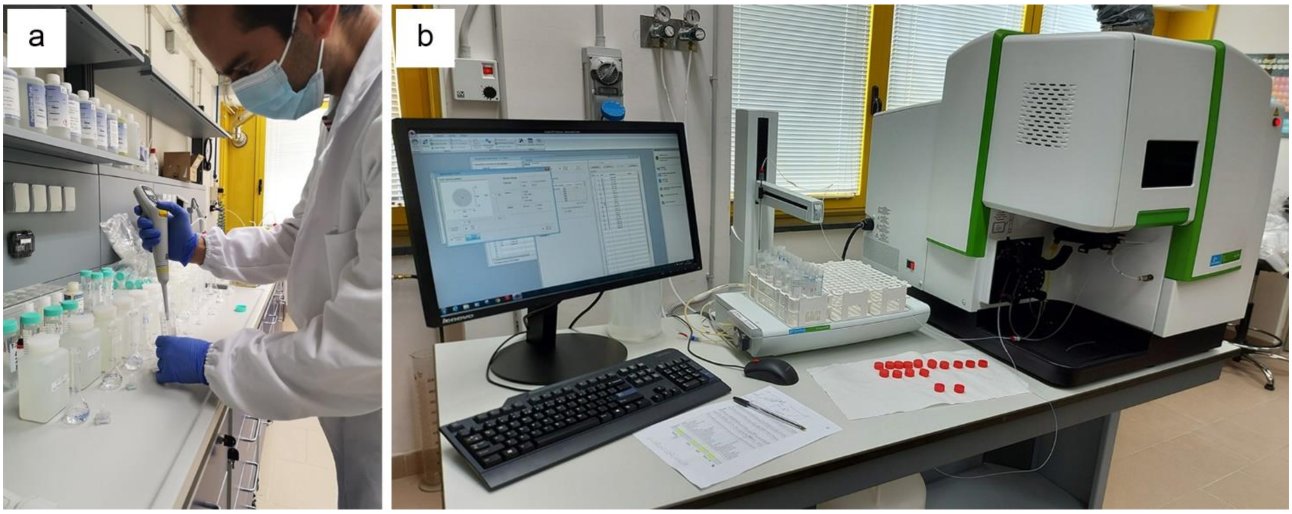


Fig. 3 (a) Sample preparation and (b) ICP-OES analysis at Department of Earth, Environmental and Resources Sciences (DiSTAR, UNINA).

Fig. 4 The measuring set-up for radon detection with AlphaGUARD PQ2000PRO: (1) AlphaGUARD; (2) degassing vessel; (3) security vessel; (4) AlphaPUMP; and (5) thermometer. For carbon dioxide measurement, both AlphaGUARD and AlphaPUMP are replaced with (6) ECOPROBE 5.

2.2. Data interpretation

During the three-year Ph.D. program, the recent publications on the advances in geochemical data analysis were studied to understand the concept and deal with the existing research gaps in the study area. In addition to the relevant online workshops, participating in a virtual Erasmus+ learning agreement and collaborating with researchers at the Center for Ecological-Noosphere Studies (Armenia) from September to December 2021 helped to learn some useful skills for (geo)chemical investigation of different environmental media. Particularly, the main focus has been:

- *compositional data analysis (CoDA)* which is getting more and more popular for analyzing geochemical data. Contrary to the classical statistical analysis, it respects the nature of geochemical data through applying an appropriate log-ratio transformation based on the research objective and transforms data from their particular geometry to the well-known Euclidean space before further analysis. Therefore, the results of the statistical analysis are not subject to spurious correlation. Scientists have been working on this line of research to propose more powerful data interpretation tools. The CoDaPack software and some R packages were applied for data treatment in this thesis.
- *spatial analysis* which is essential for geochemists worldwide because it facilitates handling a large dataset, predicting unmeasured values and visualizing the spatial distribution of elements. Classifying the maps with multifractal methods and overlaying supplementary information such as urban areas, industrial sites and volcanic centers on the maps could highlight the hot spot areas associated with a particular activity/process and identify the probable sources (natural vs. anthropogenic) of elements. The outcomes are useful for detailed geochemical studies and management practices. In this thesis, in

addition to dot maps, the inverse distance weighted (IDW) and empirical Bayesian kriging (EBK) interpolation techniques are performed and classified with multifractal analysis using QGIS and ArcGIS software.

Overall, applying a combination of the abovementioned methods in a research project could result in robust data analysis for solving a given scientific question.

3. Rationale

After the last eruption at Monte Nuovo (1538 CE), Phlegrean Fields experienced a long period of subsidence until the second half of the twentieth century when rapid ground uplift phases began: 1950-1952, 1969-1972 and 1982-1984 (Fig. 5). The uplifts were followed by subsidence, but the ground surface did not return to its pre-uplift level (Dvorak and Mastrolorenzo, 1991). Since 1984, several short bradyseismic episodes with displacements of about ≤0.1 m have been recorded in 1989, 1994, 2000, 2006, and 2012-2013 (Fig. 5; Moretti et al., 2017). During the 1970s and 1980s, risk of Phlegrean Fields volcanic eruption and the associated evacuations caused serious disruption to livelihoods mainly within Pozzuoli, the largest town inside the caldera (Barberi et al., 1984; Longo, 2019).

To understand the current situation and evaluate the risk of a volcanic eruption in Phlegrean Fields, the Civil Protection Department organizes a videoconference with the Competence Centers responsible for monitoring volcanic activity and with the Campania Region to declare the alert level (i.e. green, yellow, orange or red) every month. The alert levels represent the status of volcanic activity and highlight the period before a possible eruption. The passage from one alert level to the following one is determined based on monitoring the variations of some parameters and assessing whether the present phenomena could continue. Every six months, the Civil Protection Department and the Campania Region's civil protection structure consult with the Major Risks Commission–Volcanic Risk Sector to decide on confirming the alert levels and the operational phases (attention, early warning or alarm). Currently, the alert level of “yellow” and the operational phase of “attention” are adopted for the Phlegrean Fields. In order to enhance public awareness, an interactive map is available online (Fig. 6) to help locals understand whether they live, work or reside in the red or yellow zone. The waiting area, meeting area, means of transport, Region or Autonomous Province twinned with a given municipality and rules of conduct are also accessible for each area on the map. Therefore, it is of the utmost importance to delineate more representative precursors of volcanic eruption for enabling the authorities to manage an emergency and save lives effectively. A literature review shows that the CoDA approach was not used for hydrogeochemical studies in volcanic aquifers and this thesis tries to bridge this research gap.

The intensive anthropogenic (i.e. the industrial activity, urbanization and agricultural practices) and volcanic activities introduce potentially toxic elements to the environment. The closure of the local industrial site occurred in 1990s, but the brownfield and a part of the marine environment were listed as a polluted site of national interest for monitoring the associated environmental risk and recovering the area. More than two million people are served by the agricultural products and seafood in Phlegrean Fields and the immediate surroundings. Hence, spatial distribution, origin and the health risk of potentially toxic elements provide some information for protecting public health. There are many studies on the potentially toxic elements in soil (e.g., De Vivo et al., 2016; Petrik et al., 2018a; Thiombane et al., 2019; Guagliardi et al., 2020; Zuzolo et al., 2020) and sediment (e.g., ICRAM, 2005; Bergamin et al., 2005; Romano et al., 2004, 2009; Albanese et al., 2010; Trifuoggi et al., 2017; Armiento et al., 2020; Daliri et al., 2020; Giglioli et al., 2020), but they were not exclusive to Phlegrean Fields and the immediate surroundings or they did not consider nature of geochemical data for interpretation the results.

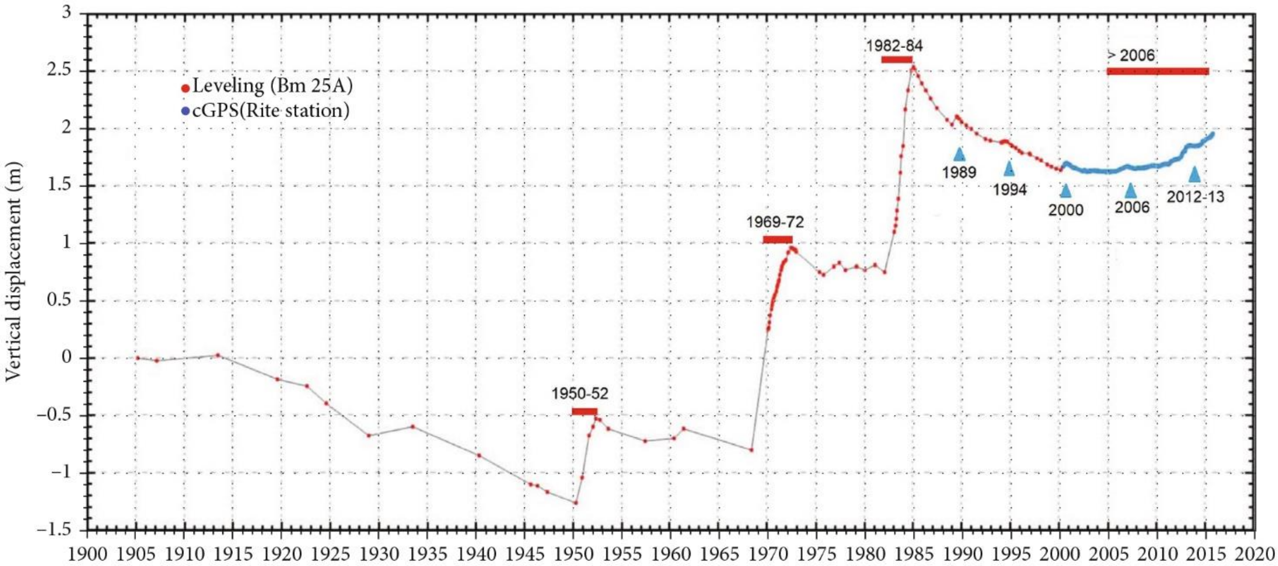


Fig. 5 Vertical displacement at Phlegrean Fields since 1905 (Moretti et al., 2017). Ground displacement was measured at Pozzuoli Porto by precision leveling (red dots: from 1905 to 2000) and at the GPS benchmark of Rione Terra (blue dots: after 2000).

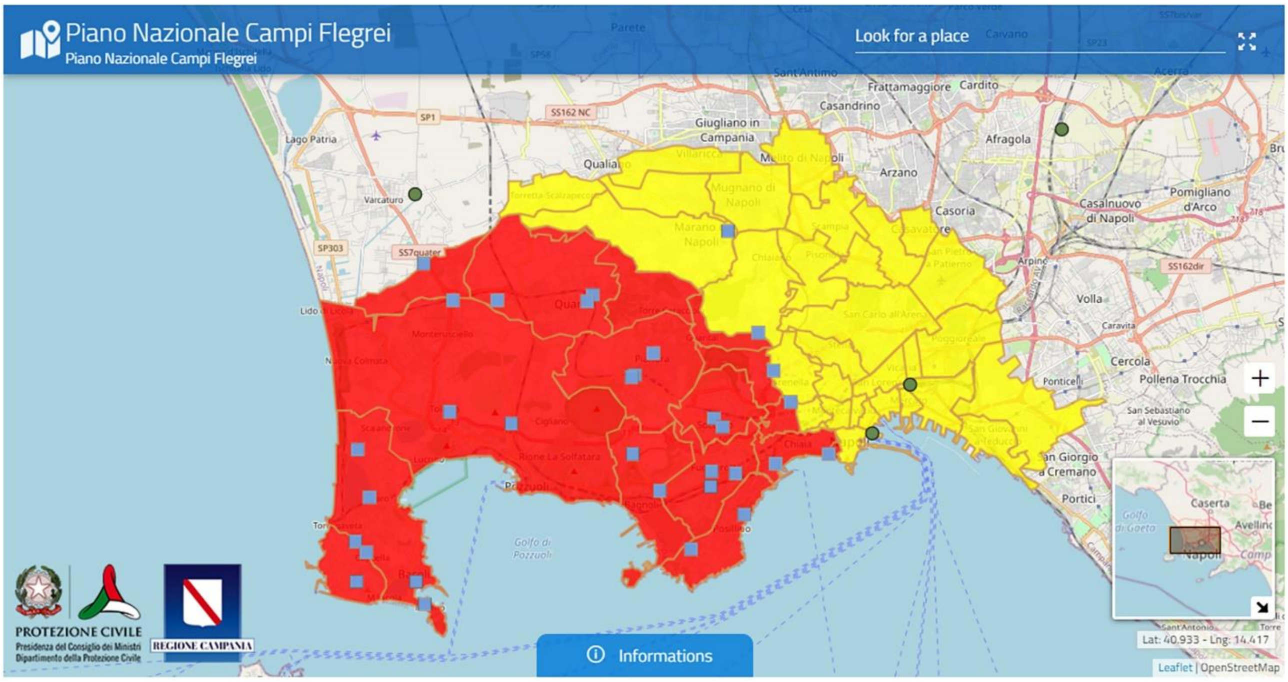


Fig. 6 The red and yellow zones for Phlegrean Fields. The figure is a screenshot of the interactive map which is available online at https://mappe.protezionecivile.gov.it/it/mappe-rischi/piano-nazionale-campi-flegrei

4. Materials in this thesis

This thesis includes seven chapters comprising an introduction, five main research chapters and a combined conclusion. Of the research chapters, the last one (i.e. chapter six) is dedicated to the scientific collaborations with other scholars for getting familiar with potentially useful research ideas in Phlegrean Fields. Subdivision of the study area for investigating each environmental compartment is presented in Fig. 7.

Chapter 2 examines groundwater hydrogeochemistry in Phlegrean Fields under two sections.

Section 2.1 explores composition of 44 groundwater samples in the Phlegrean Fields volcanic aquifer to better understand the hydrothermal system and the main processes influencing hydrogeochemical evolution of groundwater (Fig. 7). The hierarchical clustering categorized the samples into four groups: (1) bicarbonate-rich groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The composition of each group was then visualized with a range of traditional and compositional approaches. The first group is mainly derived from meteoric water with low arsenic quantities, but there is a significant thermal/seawater contribution in the second one. Interaction of the bicarbonate-rich groundwater and hydrothermal vapors leads to the sulfate-rich groundwater around Solfatara volcano. Depending on the hydrogeological factors, the mixed groundwater is observed where the three main groundwater groups undergo a mixing process. Overall, the hierarchical cluster analysis provides more information, particularly when compositional data analysis is subsequently applied to uncover the relationships between variables. This section was published in “Journal of Geochemical Exploration” as a peer-reviewed research article (Ebrahimi et al., 2022).

Section 2.2 presents a follow-up study on the dissolved radon and carbon dioxide (n = 26) in the groundwater system of Phlegrean Fields (Fig. 7). Similar to the previous section, major ions, lithium, arsenic and boron together with physicochemical parameters (i.e. pH, TDS and temperature) were also determined to use a CoDA approach for extracting the relevant information and finding the meaningful geochemical associations. Both dissolved gases share almost similar spatial distributions, but the log-ratio transformed CO_2 distinguishes bicarbonate-rich groundwater better than the raw values. Principal component analysis visualized two associations: (1) $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{K^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{HCO_3^{-}}$+$\mathrm{CO_2}$ and pH; and (2) $\mathrm{Na^{+}}$, $\mathrm{Cl^{-}}$, As, B, Li, Rn, TDS and T. It highlights that the groundwater composition is generally influenced by meteoric water (modified by the CO_2-rich magmatic gas in some cases) and hydrothermal fluid/seawater. This section is currently under review for publication as a peer-reviewed research article (Ebrahimi et al., under review).

Chapter 3 examines chemistry of tap water in Naples (Fig. 7). The Phlegrean Fields hydrothermal system negatively impacted quality of the local groundwater for human consumption and Naples tap water is supplied from the nearby aquifers and well fields. In this research, the exploited groundwater (i.e. the input of water distribution system) and tap water are monthly/seasonally collected for instrumental analysis. The seasonal changes in chemistry of the input water and the annual variations in total dissolved solids (from 200 to 1000 mg/L) of the tap waters underlines the importance of interconnections between urban water reservoirs for mixing different water types and adjusting water quality. It is complicated in populated cities like Naples with an old water distribution network and hinders setting up hydraulic and water quality models. The preliminary data visualization indicates the different chemical characteristics of some samples that are supposed to receive water from the same source. Thus, the compositional nature of chemical data was considered in hierarchical cluster analysis (HCA) to define the source of tap water in each city area. The proposed method can preliminarily divide the pipe network into unique clusters and provide an overview of the relationships between different components when representative models cannot be set up due to limited information about network characteristics. Hence, advanced water distribution simulation and management are encouraged. This chapter was published in “Environmental Science: Water Research & Technology” as a peer-reviewed research article (Ebrahimi et al., 2021).

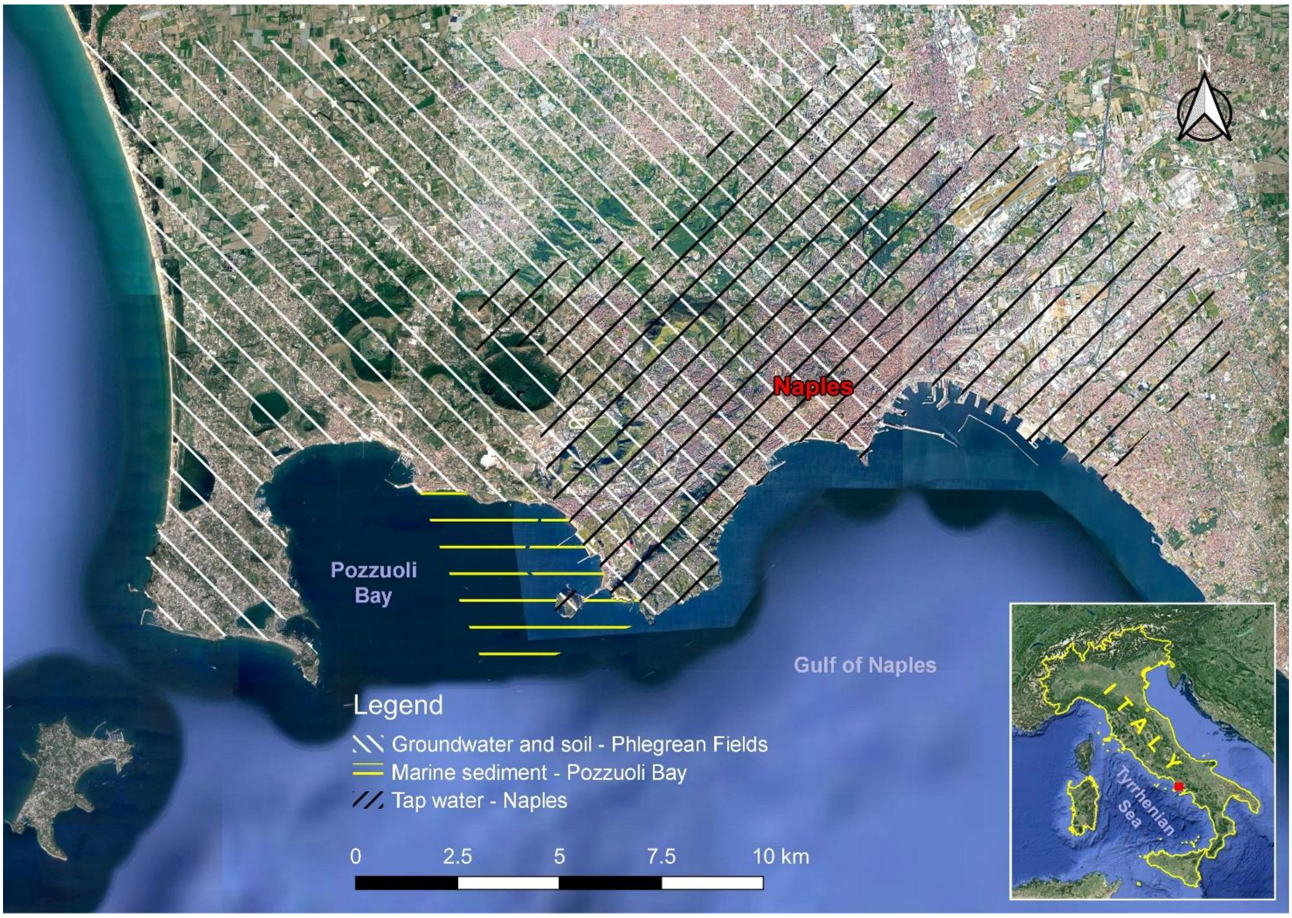


Fig. 7 A map showing subdivision of the study area into particular sectors for investigating various environmental compartments.

Chapter 4 explores sediment geochemistry in the eastern sector of Pozzuoli Bay (Fig. 7) to enlighten source of the metal(loid)s which has been a matter of debate in the last decades. In this investigation, the coastal area was divided into proximal and distal zones and compositional data analysis (CoDA) was used for data interpretation because the results are independent of the measurement unit, the selected subgroup of elements and the order of variables in the dataset. The robust principal component analysis (PCA) indicated that Hg, Cd, Cu, Pb and Zn were positively correlated with mud and organic matter in front of the former Bagnoli industrial site. Concentrations of these metals decrease along the cores and in the distal zone. On the other hand, Al, As, V, Fe, Cr, Ni and sand form an association along the coast which strengthens with increasing distance from fumaroles in the proximal zone. The results reveal that arsenic was mainly originated from the pyroclastic deposits of Phlegrean Fields and some of the seepages with a hydrothermal component, supported by the PCA in distal zone. This chapter was published in “Chemosphere” as a peer-reviewed research article (Somma et al., 2021).

Chapter 5 examines signature of the potentially toxic elements (PTEs) in soil (Fig. 7) because the intensive anthropogenic and volcanic activities might pose adverse health effects to the locals, particularly during the COVID-19 pandemic. A total of 394 topsoils (0 to 20 cm) are collected for determination of PTEs along with some indicator elements for PTE source in the fraction finer than 2 mm. Higher Pb, Zn, Cd, Cr, Hg, Ni and Sb quantities are related to the urban area in the east and southeast, but greater As, Tl, Co, Cu, Se and V levels are determined in the soils from the other parts. The Pb–Zn–Hg–Cd–Sb–Cr–Ni geochemical association probably highlights anthropogenic activities such as heavy traffic load and fossil fuel combustion in the urban area. The Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association mostly reveals the contribution of pyroclastic deposits while the Na–K–B association feasibly indicates the weathering degree, being lower in the Neapolitan Yellow Tuff caldera. The probabilistic health risk modelling for the children <6 years old shows negligible non-carcinogenic risk for exposure to Pb and Zn through soil ingestion (certainty > 95%). Pb carcinogenic risk through the ingestion route is unexpected for all age groups (CR ≤ 1E-06; certainty = 99%). However, for the inhalation pathway, there is a 90% chance of acceptable (i.e. from 1E-6 to 1E-4) Pb carcinogenic health risk for the children aged <1 year old and a 25-45% chance for those between 1 and 6 years old. This chapter is in preparation for submission to a journal as a research article (Albanese et al., in preparation).

Chapter 6 presents the collaborations with other researchers in three sections. In the first one, practicing the application of factor analysis and multifractal inverse distance weighting (MIDW) for exploring deep orebodies facilitated interpreting the groundwater, tap water, sediment and soil datasets in the previous chapters. However, the second section is a literature review performed by Pooria Ebrahimi and revised by the other author. It highlights the link between gadolinium contamination in water resources and injection of contrast agents for magnetic resonance imaging exams. Because the cities in Phlegrean Fields and the immediate surroundings are served by more than 10 magnetic resonance imaging centers that offer diagnostic imaging with contrast agent administration, quantifying the anthropogenic Gd quantities in water resources is a research gap that needs to be addressed. The last section explains bioavailability of some elements in the soils of Campania region where Phlegrean Fields is situated. Taking part in this research activity for statistical analysis and data interpretation facilitated developing the chapter on soil geochemistry in the thesis. Overall, the collaborations helped to gain more experience in working as a group member.

Section 6.1 examines the geochemistry of 404 rock and ore samples from 5 boreholes in two exploration lines of the Ashele Cu-Zn polymetallic mining area (Xinjiang, NW China) to evaluate the potential for deep and periphery prospecting. Based on the results of this research and previous studies, As, Sb, Ba, and Au are the supra-ore halo elements, Pb, Zn, Cu, and Ag (Au) are the near-ore halo elements, and Bi, Co, Mo, and Sn are the sub-ore halo elements. Factor analysis revealed exhalative-sedimentation mineralization as the first metallogenic stage followed by metamorphic-deformation and hydrothermal superimposed mineralization as the second one. The superposition area of the sub-ore halo and the supra-ore halo together with multivariate prospecting information (indicated by the primary halo zoning index, the geochemical parameter, the trilinear diagram of the primary halo and the factor score map) likely suggest occurrence of blind orebodies at higher depth of four boreholes. Some geochemical parameters also showed the multiple-periodic and multiple staged overlapping as signatures of the mineralization. This section was published in “Journal of Geochemical Exploration” as a peer-reviewed research article (Zheng et al., 2020).

Section 6.2 explains that large amounts of anthropogenic Gd have been released into surface water due to the unmetabolized excretion of Gd chelates administered in magnetic resonance imaging exams and the ineffective removal process in wastewater treatment plants. Number of gadolinium-based contrast agent (Gd-CA) administrations is expected to continue growing and introduce greater anthropogenic Gd quantities in water resources. The feasibility of Gd retention after repeated Gd-CA administrations and even potentially fatal diseases have recently arisen severe health concerns. It shows the need to investigate the probable adverse health effects of the currently marketed Gd-CAs and to modify the approach in using Gd contrast media in daily practices. The employment of enhanced wastewater treatment processes for removing the stable contrast agents, the evaluation of Gd chelate ecotoxicity and the assessment of human exposure levels deserve more attention as well. On the other hand, anthropogenic Gd microcontamination presents the opportunity to assess surface water–groundwater interactions and trace the fate of other microcontaminants associated with treated wastewater plumes. This section was published in “Geosciences” as a review article (Ebrahimi and Barbieri, 2019).

Section 6.3 explores the pseudo-total concentrations and the bioavailable fractions of Al, Ca, Cu, K, Mg and Tl in 1993 topsoil samples from Campania (southern Italy). The datasets were compared through geochemical mapping and statistical analysis to highlight the spatial patterns and the key influencing factors. The medians for the bioavailable percentage of elements followed the order of Ca > K >> Mg ≃ Tl >> Cu >> Al. A multiple linear regression model was applied to the data, including grain size and organic matter content, and it was revealed that the pseudo-total concentration and grain size play an important role in estimating the bioavailable fraction of most elements. Organic matter is a significant input variable for predicting the bioavailable Ca and K contents. In the future investigations, it is advisable to consider more parameters that regulate bioavailability of elements to achieve more representative estimations in the models. This section was submitted to a journal as a research article (Guarino et al., submitted).

5. Publications, manuscripts and presentations/posters

The peer-reviewed publications in line with the Ph.D. project (the third, fifth and sixth) and the other scientific collaborations include:

- Abbasi, S., Sheikh Fakhradini, S., Jaafarzadeh, N., Ebrahimi, P., & Ashayeri, S. Y. (2021). Eutrophication and sediment–water exchange of total petroleum hydrocarbons and heavy metals of Hashilan wetland, a national heritage in NW Iran. Environmental Science and Pollution Research, 1-19. https://doi.org/10.1007/s11356-021-17937-x.
- Abbasi, S., Yavar Ashayeri, S., Jafarzadeh, N., Fakhradini, S. S., Alirezazadeh, M., Ebrahimi, P., ... & Pashaei, R. (2021). Hydrological and hydrogeological characteristics and environmental assessment of Hashilan Wetland, a national heritage in NW Iran. Ecohydrology & Hydrobiology. https://doi.org/10.1016/j.ecohyd.2021.08.014.
- 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. https://doi.org/10.1039/D0EW01123A.

- Ebrahimi, P., & Barbieri, M. (2019). Gadolinium as an emerging microcontaminant in water resources: threats and opportunities. Geosciences, 9(2), 93. https://doi.org/10.3390/geosciences9020093.
- Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., ... & Albanese, S. (2022). Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy). Journal of Geochemical Exploration, 106922. https://doi.org/10.1016/j.gexplo.2021.106922.
- 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. https://doi.org/10.1016/j.chemosphere.2021.129955.
- Zheng, C., Luo, X., Wen, M., Ebrahimi, P., Liu, P., Liu, G., ... & Albanese, S. (2020). Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China. Journal of Geochemical Exploration, 213, 106509. https://doi.org/10.1016/j.gexplo.2020.106509.

The prepared manuscripts in line with the Ph.D. project (the first two) and in collaboration with other colleagues are:

- Albanese, S., Ebrahimi, P., Aruta, A., Cicchella, D., De Vivo, B., & Lima, A. Potentially toxic elements in the soils of Phlegrean Fields (south Italy) and the immediate surroundings: Spatial distribution, origin and probabilistic human health risk – *In Preparation*
- Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Cicchella, D., & Albanese, S. The relationship between dissolved gases (radon and carbon dioxide) and other geochemical signatures in Phlegraean Fields volcanic aquifer: a follow-up study – *Under Review*
- Guarino, A., Albanese, S., Cicchella, D., Ebrahimi, P., Dominech, S., Allocca, C., Romano, N., De Vivo, B., & Lima, A. Selected major and potentially toxic elements bioavailability in agricultural soils of Campania region (Italy): The spatial patterns and influencing factors – *Submitted*

In addition, the results of this thesis (i.e. the fifth and sixth items below) and other research projects were presented in several (inter)national conferences:

- Aruta, A., Guarino, A., Ebrahimi, P., Dominech, S., Belyaeva, O., Tepanosyan, G. & Albanese, S. (2021). Low-level ionizing radiation and associated risk in an urban environment: the importance of both paving and building materials, *EGU General Assembly 2021*, online, 19-30 Apr, EGU21-5566. https://doi.org/10.5194/egusphere-egu21-5566.
- Aruta, A., Guarino, A., Ebrahimi, P., Dominech, S., Belyaeva, O., Tepanosyan, G. & Albanese, S. (2020). Radiological risk induced by road paving and building materials in the city of Naples, south Italy. In: S. Fiore (Editor). *I International Meeting of GeoHealth Scientists* - GHC 2020 BARI, Italy. Digilabs Pub., Bari, Italy, pp. 51.
- Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S. & Albanese S. (2020). Potentially toxic elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (Southern Italy). In: S. Fiore (Editor). *I International Meeting of GeoHealth Scientists* - GHC 2020 BARI, Italy. Digilabs Pub., Bari, Italy, pp. 16.
- 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. https://doi.org/10.7185/gold2021.6064.
- Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., Capecchiacci, F., Carandente, A., Minopoli, C., Santi, A. & Albanese, S. (2021). Fluoride in Campi Flegrei volcanic aquifer, south Italy: A comparison between water and rock composition, *BE GEOSCIENTISTS 2021*, 7-10 October, Naples, Italy. https://doi.org/10.3301/ABSGI.2021.04.

- Ebrahimi, P., Somma, R., Troise, C., De Natale, G., Guarino, A., Cicchella, D., & Albanese, S. (2021). Characterizing source of arsenic in the sediments of Pozzuoli Bay (south Italy) using compositional data analysis (CoDA). *Goldschmidt2021• Virtual*• 4-9 July. https://doi.org/10.7185/gold2021.6152.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S. & Albanese, S. (2020). The "Triangle of Death": A Case Study from Campania Region (Italy). *Goldschmidt2020• Virtual*• 21-26 June. https://doi.org/10.46427/gold2020.897.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S., Lima, A., Cicchella, D. & Albanese, S. (2021). Radon fluxes estimate from geochemical data and gamma radiation in Campania region (Italy), *BE GEOSCIENTISTS 2021*, 7-10 October, Naples, Italy. https://doi.org/10.3301/ABSGI.2021.04.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S., Lima, A., De Vivo, B., ... & Albanese, S. (2021). Organochlorine pesticides in the soils of the Acerra plain: concentration and distribution of DDT isomers and metabolites, *EGU General Assembly 2021*, online, 19-30 Apr, EGU21-5739. https://doi.org/10.5194/egusphere-egu21-5739.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S., Lima, A., De Vivo, B., Qi, S., & Albanese, S. (2020). Potentially harmful elements and polycyclic aromatic hydrocarbons in the soils of Acerra, southern Italy. In: S. Fiore (Editor). *I International Meeting of GeoHealth Scientists* - GHC 2020 BARI, Italy. Digilabs Pub., Bari, Italy, pp. 10.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S., Zuzolo, D., Lima, A., Cicchella, D. & Albanese, S. (2020). Pedogenic radon fluxes predictions from geochemical data and gamma ray: The Campania region experiment. *I International Meeting of GeoHealth Scientists* - GHC 2020 BARI, Italy. Digilabs Pub., Bari, Italy, pp. 51.
- Zuzolo, D., Albanese, S., Iannone, A., Ebrahimi, P., Melito, R., Guarino, A., Aruta, A. & Cicchella, D. (2020). Reconnaissance soil gas radon survey in Campania region (Italy): Preliminary results. In: S. Fiore (Editor). *I International Meeting of GeoHealth Scientists* - GHC 2020 BARI, Italy. Digilabs Pub., Bari, Italy, pp. 51.

After participating in the I International Meeting of GeoHealth Scientists in 2020, the following conference proceedings were also prepared in collaboration with the group members:

- Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S. & Albanese S. (2020). Potentially toxic elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (Southern Italy): a focus on cadmium. *Proceedings of I International Meeting of GeoHealth Scientists* - GHC 2020, BARI (Italy), *Proscience*, 7, 35-40. https://www.doi.org/10.14644/ghc2020.006.
- Guarino, A., Aruta, A., Ebrahimi, P., Dominech, S., Lima, A., De Vivo, B., Qi, S., & Albanese, S. (2020). Potentially harmful elements and polycyclic aromatic hydrocarbons in the soils of Acerra, southern Italy. *Proceedings of I International Meeting of GeoHealth Scientists* - GHC 2020, BARI (Italy), *Proscience*, 7, 41-46. https://www.doi.org/10.14644/ghc2020.007.

CHAPTER 2

Groundwater Hydrogeochemistry in Phlegrean Fields

Section 2.1 was published in “Journal of Geochemical Exploration” as a Research Article:

> Ebrahimi, P. *, Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., ... & Albanese, S. (2022). Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy). Journal of Geochemical Exploration, 106922. https://doi.org/10.1016/j.gexplo.2021.106922.
>
> * Corresponding author: pooria.ebrahimi@unina.it; pooria.ebrahimi@gmail.com

A summary of the results was also presented in BE GEOSCIENTISTS 2021 (Ebrahimi et al., 2021).

Section 2.2 is currently under review in a journal as a Research Article:

> Ebrahimi, P. *, Guarino, A., Allocca, V., Caliro, S., Cicchella, D., & Albanese, S. The relationship between dissolved gases (radon and carbon dioxide) and other geochemical signatures in Phlegraean Fields volcanic aquifer: a follow-up study.
>
> * Corresponding author: pooria.ebrahimi@unina.it; pooria.ebrahimi@gmail.com

Section 2.1:
Hydrogeochemical Evolution of Groundwater

Abstract

Comprehensive hydrogeochemical studies have been conducted in the Phlegrean Fields volcanic aquifer since late 20th century due to the volcanic unrest. In the last decade, groundwater samples were grouped based on the dominant anion species (i.e. bicarbonate, sulfate and chloride) to explain the general hydrogeochemical processes. In this section, 44 groundwater samples are collected from Phlegrean Fields 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 groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The first group (As = 158.2 ± 169 µg/L, electric conductivity = 1732.1 ± 1086 µS/cm and temperature = 25.6 ± 8 °C) is mainly derived from poor-arsenic meteoric water, but there is significant thermal/seawater contribution in the second one (As = 1457.8 ± 2210 µg/L, electric conductivity = 20118.3 ± 11139 µS/cm and temperature = 37.1 ± 20 °C). Interaction of the bicarbonate-rich groundwater and hydrothermal vapors gives rise to the sulfate-rich groundwater (As = 847.2 ± 679 µg/L, electric conductivity = 3940.0 ± 540 µS/cm and temperature = 82.8 ± 3 °C) around Solfatara volcano. The mixed groundwater (As = 451.4 ± 388 µg/L, electric conductivity = 4482.9 ± 4027 µS/cm and temperature = 37.1 ± 16 °C) is observed where the three main groundwater groups undergo a mixing process, depending on the hydrogeology of the volcanic aquifer. Contrary to the bicarbonate- and sulfate-rich groundwater, the chlorine-rich and mixed groundwater generally occurs at low piezometric levels (approximately <1 m above sea level) near the coastline. The hierarchical cluster analysis provides 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 and identifying promising precursors of volcanic eruption.

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.

Phlegrean Fields are the active volcanic areas, well-known for the striking ground movements or bradyseism, in which thermal energy is mainly released through diffuse degassing (Chiodini et al., 2001). Groundwater hydrogeochemistry was extensively investigated in the Phlegrean Fields 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, Phlegrean Fields 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 Phlegrean Fields 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 explains only a part of the whole composition. Karl Pearson introduced the problem of spurious correlation in this type of data (Pearson, 1897) 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, 2022; 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 research 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 Phlegrean Fields 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 CE. 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 Phlegrean Fields 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 the 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 m 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). The variations in pH, ammonia and/or $\mathrm{SO_4^{2-}/Cl^-}$ 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).

Little is known about the hydrologic system of Phlegrean Fields at the southwestern boundary of Campania Plain. 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. A 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).

Various origins were proposed for the thermal waters in Phlegrean Fields: (1) Seawater or brines probably exist at the 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 Phlegrean Fields (Caprarelli et al., 1997). High As concentrations in the Na–Cl brines originate 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 groundwater during their ascent towards the surface and give rise to mixed water with intermediate to low arsenic content (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 the 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 the As-bearing hydrothermal steam results in intermediate to high As in the 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 endmembers in Phlegrean Fields: (1) alkali-chloride, highly saline water (Stufe di Nerone well and Stufe di Nerone spring); (2) alkalibicarbonate, less saline water (Agnano spring and the water from Quarto plain); (3) alkali-bicarbonate-sulfate water (a sampling station in Quarto plain); and (4) alkaline earth-sulfate, acidic water (Pisciarelli Grande and Pisciarelli Piccola). They proposed that the groundwater from Pozzanghera Agnano Terme, Hotel Tennis, Terme Puteolane and Damiani is mixture of the endmembers. However, Aiuppa et al. (2006) used the percent meq/L of $\mathrm{HCO_3^-}$, $\mathrm{SO_4^{2-}}$ and $\mathrm{Cl^-}$, defined the prevalent anion species and grouped 64 samples in bicarbonate-, chlorine- and sulfate-rich groundwater. The latter classification can be considered as the generalized version of the former, being beneficial when investigating a large number of samples.

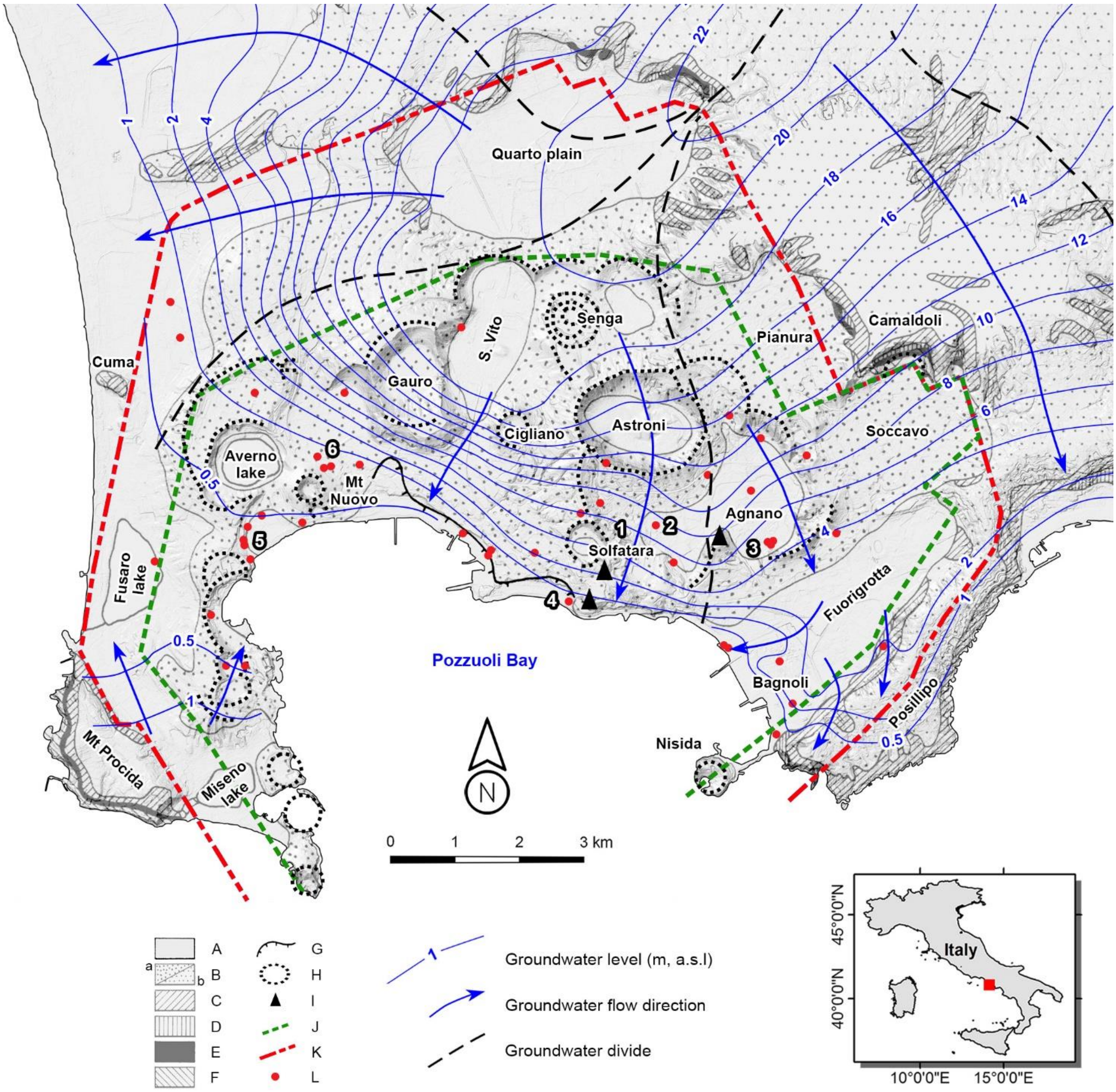


Fig. 1 The simplified geological map of Phlegrean Fields (after Valentino et al., 1999). The groundwater level, groundwater flow direction, groundwater divide (after De Vita et al., 2018) together with location of (1) Pisciarelli, (2) Hotel Tennis, (3) Agnano Terme, (4) Terme Puteolane, (5) Stufe di 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); and (L) sampling points.

3. Materials and methods

3.1. Sampling and chemical analysis

The majority of published hydrogeochemical data in Phlegrean Fields 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 section, a total of 44 water samples are collected in May 2019 from wells and springs mostly located in the NYT caldera (Fig. 1; Table 1). Alkalinity, electric conductivity (EC), pH and temperature are measured with portable instruments in situ. 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 study. The R package “stats” (R Core Team, 2021) is utilized to perform hierarchical clustering.

Table 1 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

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 Phlegrean Fields (Aiuppa et al., 2006):

$$\mathrm{Cl^-\ ratio} = \frac{\mathrm{Cl^-}}{\mathrm{HCO_3^- + SO_4^{2-} + Cl^-}} \tag{1}$$

$$\mathrm{SO_4^{2-}\ ratio} = \frac{\mathrm{SO_4^{2-}}}{\mathrm{HCO_3^- + SO_4^{2-} + Cl^-}} \tag{2}$$

and

$$\mathrm{HCO_3^-\ ratio} = \frac{\mathrm{HCO_3^-}}{\mathrm{HCO_3^- + SO_4^{2-} + Cl^-}} \tag{3}$$

where all values are in mg/L. Presence of outliers and variance of the abovementioned 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).

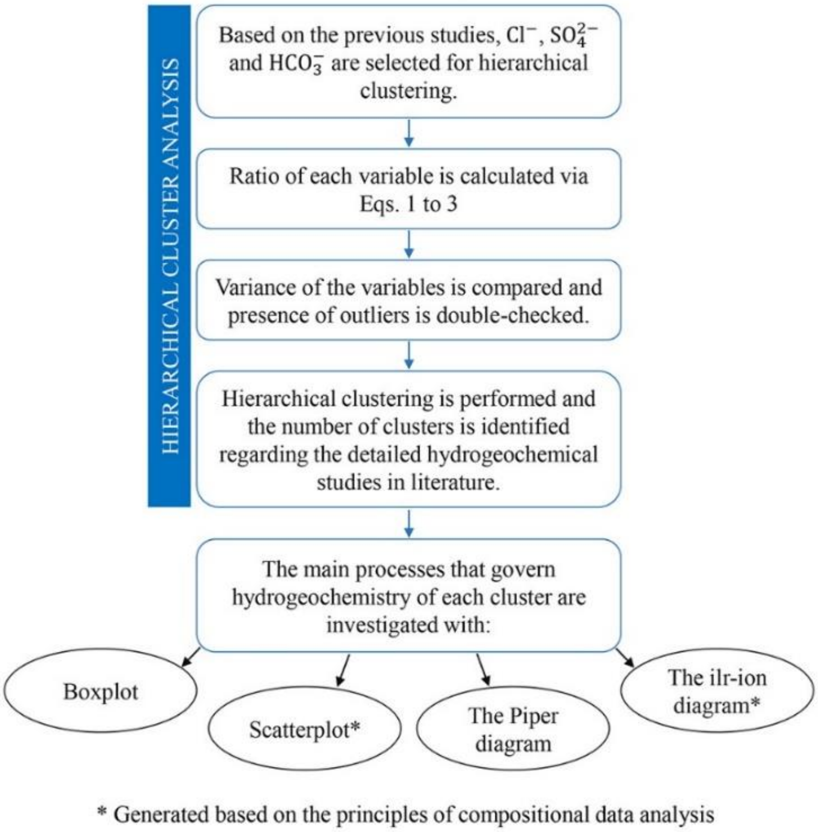


Fig. 2 The flowchart demonstrating the main data treatment and visualization steps.

3.2.2. Compositional data analysis

The 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, 1897; 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 the waters mainly composed of $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{SO_4^{2-}}$ and $\mathrm{Cl^-}$ 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 abovementioned 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:

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

$$z_2 = \frac{1}{\sqrt{2}} \ln \frac{[\mathrm{Ca^{2+}}]}{[\mathrm{Mg^{2+}}]} \tag{5}$$

$$z_3 = \sqrt{\frac{2}{3}} \ln \frac{\sqrt{[\mathrm{Cl^-}][\mathrm{SO_4^{2-}}]}}{[\mathrm{HCO_3^- + CO_3^{2-}}]} \tag{6}$$

and

$$z_4 = \frac{1}{\sqrt{2}} \ln \frac{[\mathrm{Cl^-}]}{[\mathrm{SO_4^{2-}}]} \tag{7}$$

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 (e.g., Somma et al., 2021). The R package “robCompositions” (Filzmoser et al., 2018) is used for computing the symmetric coordinates in Figs. 9 and 10. The interested readers are referred to Kynčlová et al. (2017) for the calculation steps. 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 the Na–Cl brines and seawater whilst bicarbonate is characteristic of meteoric water and the CO_2-rich hydrothermal water. Generating the following balance can describe the sources of Na^+ and Cl^- in comparison with those of HCO_3^-:

$$[\mathrm{Na}^+,\mathrm{Cl}^-|\mathrm{HCO}_3^-] = \sqrt{\frac{2}{3}}\ln\frac{\sqrt{[\mathrm{Na}^+][\mathrm{Cl}^-]}}{[\mathrm{HCO}_3^-]} \tag{8}$$

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

$$[\mathrm{Na}^+|\mathrm{Cl}^-] = \frac{1}{\sqrt{2}}\ln\frac{[\mathrm{Na}^+]}{[\mathrm{Cl}^-]} \tag{9}$$

where all concentrations are in mg/L. The R code for generating a scatterplot based on Eqs. (8) and (9) is available below:

```r
library(ggplot2) # Load the package for generating a scatterplot
# Set the working directory and import the data (please notice that name of the csv file might be
# something different). The columns “Na”, “Cl” and “HCO3” must represent the sodium, chlorine and
# bicarbonate (in mg/L), respectively. The “Group” column refers to the groups of groundwater samples.
dat = read.csv("data.csv")
# Calculate the ilr coordinates based on the equations 8 and 9
dat$NaCl_HCO3 = sqrt(2/3)*log(sqrt(dat$Na*dat$Cl)/dat$HCO3)
dat$Na_Cl = 1/sqrt(2)*log(dat$Na/dat$Cl)
# Generate the scatterplot
ggplot(dat, aes(x = Na_Cl, y = NaCl_HCO3, shape = Group, color = Group)) +
  geom_point(size = 2.7) +
  geom_vline(xintercept = 0, lty = "dotted", color = "gray") + # Add vertical dotted line
  geom_hline(yintercept = 0, lty = "dotted", color = "gray") + # Add horizontal dotted line
  labs(x = "[Na|Cl]", y = "[Na,Cl|HCO3]") + # Add x and y axis labels
  theme_bw() +
```

```
# Select point colors manually
scale_color_manual(name = "Legend",
                   breaks = c("C1 (BGW)", "C2 (MGW)", "C3 (ClGW)", "C4 (SGW)", "LSW", "GEW"),
                   values = c("#00cc00", "#0000ff", "#ff0000", "#ff9900", "#666666", "#ff0000"),
                   labels = c(expression(C1["BGW"]), expression(C2["MGW"]),
                              expression(C3["ClGW"]), expression(C4["SGW"]), "LSW", "GEW")) +
# Select shapes manually
scale_shape_manual(name = "Legend",
                   breaks = c("C1 (BGW)", "C2 (MGW)", "C3 (ClGW)", "C4 (SGW)", "LSW", "GEW"),
                   values = c(0, 1, 2, 5, 15, 8),
                   labels = c(expression(C1["BGW"]), expression(C2["MGW"]),
                              expression(C3["ClGW"]), expression(C4["SGW"]), "LSW", "GEW")) +
theme(legend.position = "bottom", # Determine legend position
      panel.grid = element_blank(), # Remove the grids
      legend.title = element_text(size = 14, color = "black"), # Adjust size and color of legend title
      legend.text = element_text(size = 12, color = "black"), # Adjust size and color of legend text
      axis.title = element_text(size = 14, color = "black", face = "bold"), # Adjust size and color of axis title
      axis.text = element_text(size = 14, color = "black")) # Adjust size and color of axis text
```

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 2). To avoid confusion, the abbreviations (i.e. BGW, ClGW and SGW which stand for bicarbonate-, chlorine- and sulfate-rich groundwater, respectively) are exclusively used for discussing the dominant anion groups. However, the HCA groups are mainly referred to as bicarbonate-rich, mixed, chlorine-rich and sulfate-rich groundwater ($C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$, respectively).

4.1. An overview of groundwater composition

Na–Cl and Na–HCO_3 are the main groundwater types in Phlegrean Fields (Table 2). However, Na–SO_4 is exclusive to the groundwater around Solfatara crater with the highest temperatures (>79 °C). On average, Na–Cl groundwater is 10 °C warmer than Na–HCO_3 groundwater. 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 to 8.6 in Na–HCO_3 groundwater revealing the acidic to alkaline nature of the groundwaters under investigation. Nonetheless, electric conductivity indicates a reverse trend, decreasing from Na–Cl groundwater to Na–HCO_3 groundwater (Table 2).

Regarding the dominant anion groups in Table 2, BGW (T = 16.7 to 53.4 °C and EC = 350 to 6340 µS/cm) mostly represents the meteoric-derived water which can be supported by the isotopic investigations of Caprarelli et al. (1997). Mixing this groundwater 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 °C and EC = 795 to 37000 µS/cm). On the other hand, the interaction between BGW and the vapor phase forms SGW with high temperature (>79 °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 H_2S by atmospheric gases results in elevated sulfate concentration in the groundwater around Solfatara crater.

4.2. Detection of homogeneous groundwater 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 (Fig. 3a to c). However, computing the anion ratios alleviates these data problems for applying clustering algorithm (Fig. 3d to f). The collected samples can be categorized in the bicarbonate-rich, sulfate-rich, chlorine-rich and mixed groundwater based on Fig. 4. 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 4 and Table 3). As the groundwater body flows towards the coast, the mixed and chlorine-rich groundwater become dominant. It is worth mentioning that the water from Agnano spring (W40 in Tables 1 and 2) 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 anion groups, hierarchical cluster analysis divides ClGW into (Fig. 4):

(1) *mixed groundwater ($C2_{MGW}$)*. Valentino and Stanzione (2003) introduced the water samples collected in Pozzanghera Agnano Terme, Hotel Tennis, Terme Puteolane and Damiani as a mixture of the endmembers. 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. 4), but HCA results express that it underwent a mixing process which is consistent with the measurements in the late 20th century (Table 4).

(2) *chlorine-rich groundwater ($C3_{ClGW}$)*. The groundwater around Stufe di Nerone well and Stufe di Nerone spring, and the groundwater 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 the groundwater from Stufe di Nerone well and Stufe di Nerone spring also suggested 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 groundwater 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. 4a, the statistically significant difference in average Cl^- and HCO_3^- ratios between them reveals their effective clustering (Fig. 5). It is noteworthy that HCA considers the multivariate data space through calculating the distances between observations. Significance of difference in average of the ratios between other combinations of clusters is not assessed due to their few observations and limitation of the bootstrapping approach. The sulfate-rich groundwater is the next group that merges these two at a higher level of the dendrogram (Fig. 4a), probably showing a relatively stronger fingerprint of meteoric water in composition of groundwater in the study area, except for the chlorine-rich groundwater.

Chemical signature of the four groups of HCA versus the three groups of dominant anion species are compared in Fig. 6. It is evident that the percentages of bicarbonate and sulfate (left panels in Fig. 6a and c) decrease from the meteoric endmember ($C1_{BGW}$) to the thermal/seawater endmember ($C3_{ClGW}$) although chloride percentage, arsenic, boron, lithium, electric conductivity and temperature (left panels in Fig. 6b, d, e, f, g and h, respectively) show a reverse trend. Values of the selected parameters in the mixed groundwater range between the corresponding values in bicarbonate- and chlorine-rich groundwater which confirm their hybrid nature. Compared to the bicarbonate-rich groundwater (As = 158.2 ± 169 µg/L, B = 2875.3 ± 3973 µg/L and EC = 1732.1 ± 1086 µS/cm), the average arsenic, boron and electric conductivity values are almost double in the mixed groundwater (As = 451.4 ± 388 µg/L, B = 5955.1 ± 5679 µg/L and EC = 4482.9 ± 4027 µS/cm) and one order of magnitude greater in the chlorine-rich groundwater (As = 1457.8 ± 2210 µg/L, B = 21787.4 ± 25278 µg/L and EC = 20118.3 ± 11138 µS/cm). The bicarbonate-rich groundwater is the coldest (25.6 ± 8 °C), but the chlorine-rich groundwater is the most acidic (pH = 6.6 ± 0.4; Table 2). The sulfate-rich groundwater has the lowest Cl^{-} (%), but the highest SO_4^{2-} (%) and temperature (Fig. 6b, c and h, respectively). The average arsenic, boron and electric conductivity values in this group (847.2 ± 679 µg/L, 14678.8 ± 13221 µg/L and 3940.0 ± 540 µS/cm, respectively) lie between those of bicarbonate- and chlorine-rich groundwater. A part of this information cannot be obtained from categorizing the samples based on the dominant anion species (i.e. the right panels in Fig. 6a to h).

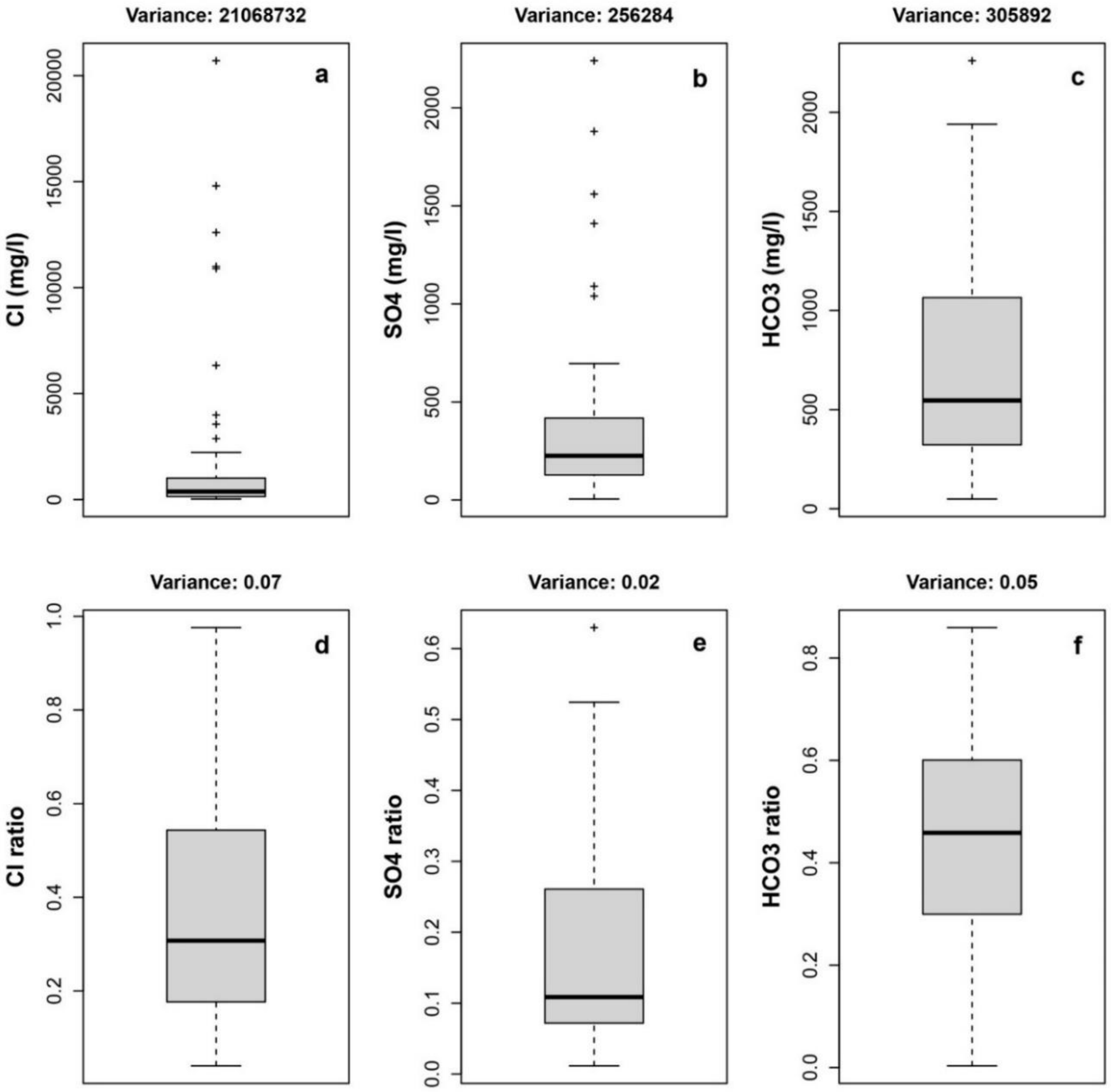


Fig. 3 Boxplots of major anions (a to c) and their ratios (d to f) calculated via equations 1 to 3. Variance of each variable is also represented above the corresponding chart.

Table 2 The measured values of ions, elements and physicochemical parameters in Phlegrean Fields groundwater samples.

ID	**Groundwater 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	36.0	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	20.0	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	56.0	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	67.0	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	75.0	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	77.0	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	32.0	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	21.0	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

Table 2 (continued)

ID	**Groundwater 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**
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	26.0	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

<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 Li, As and B (µg/L), electric conductivity (EC; µS/cm), temperature (T; ºC) and pH (unitless). [a] The groups are based on the dominant anion species in meq/L: SGW, ClGW and BGW stand for sulfate-, chlorine- and bicarbonate-rich groundwater, respectively; [b] The groups are based on hierarchical cluster analysis: C1$_{BGW}$, C2$_{MGW}$, C3$_{ClGW}$ and C4$_{SGW}$ denote the bicarbonate-rich, mixed, chlorine-rich and sulfate-rich groundwater, respectively; [c] Guglielminetti (1986); [d] Aiuppa et al. (2006)

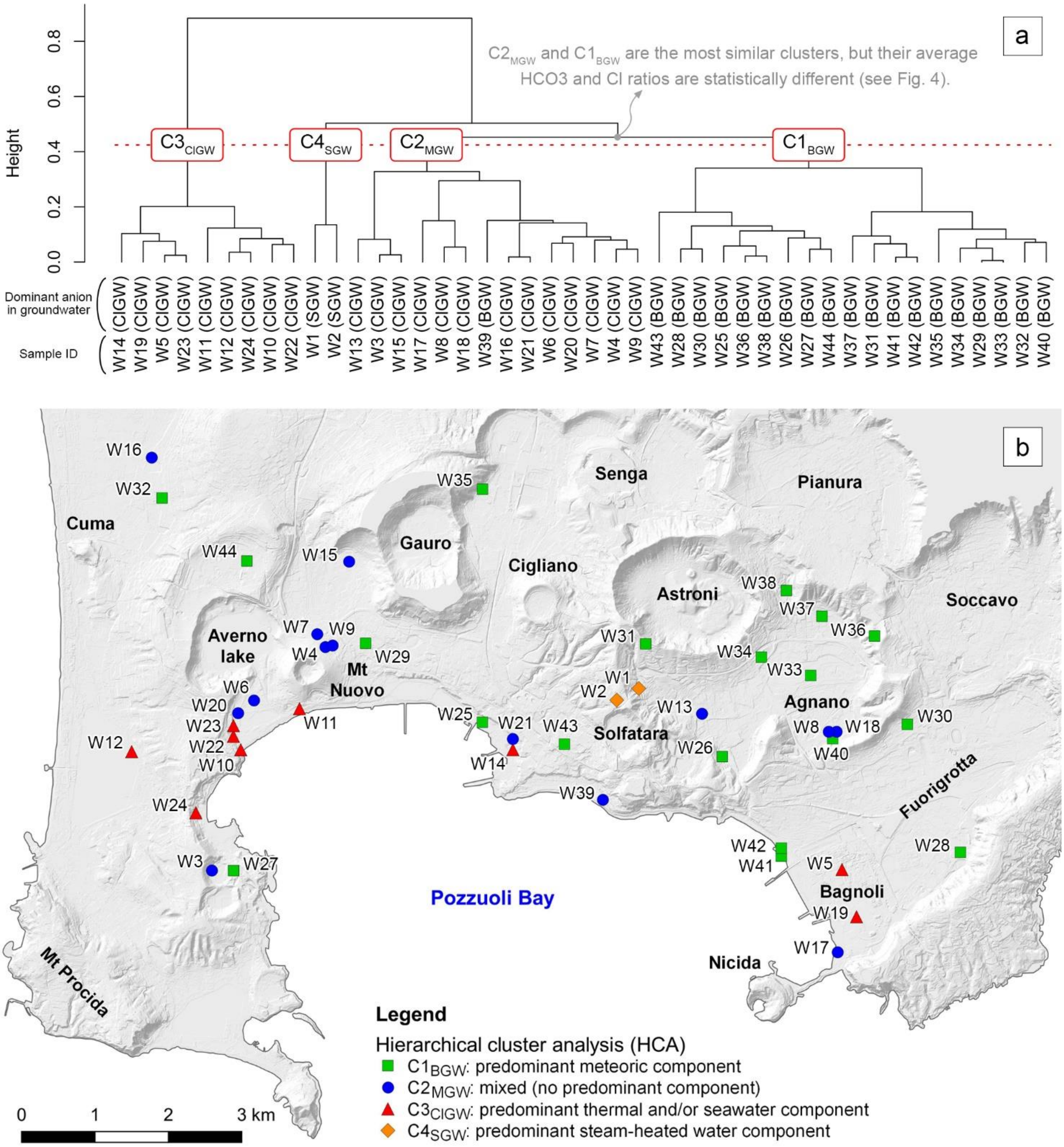


Fig. 4 (a) The dendrogram of hierarchical cluster analysis; and (b) spatial distribution of each group of groundwater samples. The height on the vertical axis of the dendrogram states the distance measure, indicating higher similarity at a lower value. On the horizontal axis of the dendrogram, BGW (bicarbonate-rich groundwater), ClGW (chlorine-rich groundwater) and SGW (sulfate-rich groundwater) refer to the dominant anion group. C1$_{BGW}$: bicarbonate-rich groundwater; C2$_{MGW}$: mixed groundwater; C3$_{ClGW}$: chlorine-rich groundwater; and C4$_{SGW}$: sulfate-rich groundwater.

Table 3 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 groundwater, 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 4 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)

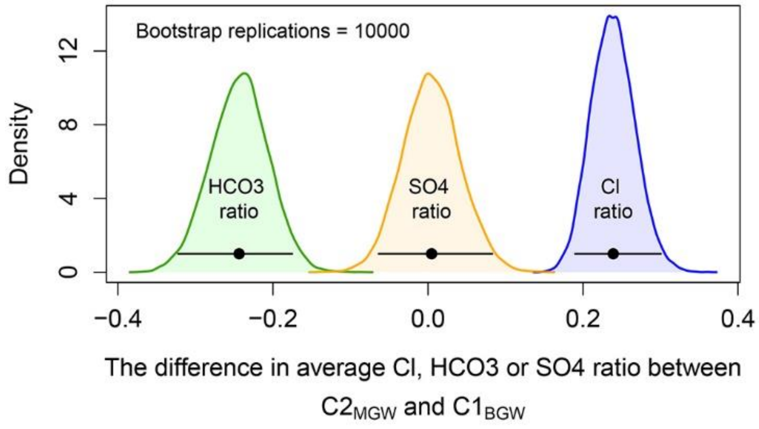


Fig. 5 Density plot of the difference in average HCO_3^-, SO_4^{2-} or Cl^- ratio between the mixed ($C2_{MGW}$) and bicarbonate-rich ($C1_{BGW}$) groundwater. The dots show the values obtained from the dataset in the present study and the horizontal line indicates the uncertainty based on the 95% confidence 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.

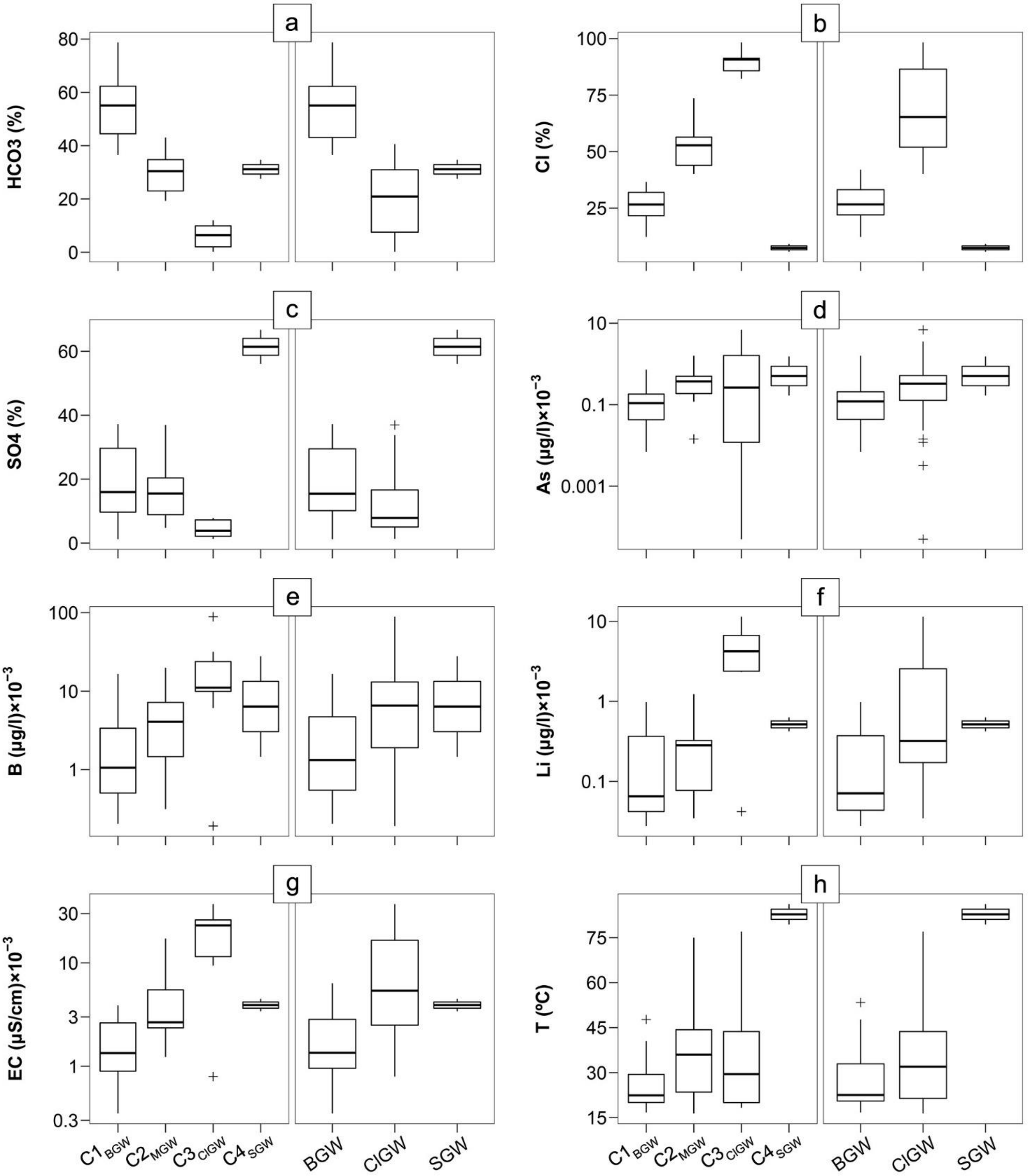


Fig. 6 Variation of selected parameters in the HCA groups and dominant anion groups (the left and right panels in each paired graph, respectively). To generate panels a to c, the following calculations are made using bicarbonate, chloride and sulfate in meq/L: $\mathrm{HCO_3^-} = (\mathrm{HCO_3^-}/(\mathrm{HCO_3^-} + \mathrm{Cl^-} + \mathrm{SO_4^{2-}})) \times 100$; $\mathrm{Cl^-} = (\mathrm{Cl^-}/(\mathrm{HCO_3^-} + \mathrm{Cl^-} + \mathrm{SO_4^{2-}})) \times 100$; 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 and the abbreviations on the horizontal axes are the same with Fig. 4.

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. 7) and ilr-ion plot (Fig. 8). 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 isometric log-ratios in Fig. 8, the chlorine-rich groundwater is not condensed in the corners of the ilr-ion plot and their chemical variation can be assessed in detail.

Excluding the sulfate-rich groundwater, the positive correlation between the samples is better demonstrated in Fig. 8b. The bicarbonate-rich groundwater shows 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 groundwater confirming higher contribution of seawater and/or the Na–Cl brine. It is noteworthy that the greatest variation of $[Cl^-, SO_4^{2-}|HCO_3^- + CO_3^{2-}]$ is observed in the chlorine-rich groundwater whilst significant changes of $[Cl^-|SO_4^{2-}]$ is evident in the bicarbonate-rich groundwater (Fig. 8b). It might show the complex situation in the study area which cannot be detected in the original Piper diagram (Fig. 7c).

Although it is difficult to distinguish all groups in cation trilinear plot of the Piper diagram (Fig. 7b) due to significant data clustering, the data points are scattered in the ilr-ion plot (Fig. 8c) which allows to decipher the underlying hydrogeochemical interactions. The Ca^{2+}/Mg^{2+} log-ratio indicates the smallest value in seawater and increases with decreasing Ca^{2+} and Mg^{2+} versus $Na^+ + K^+$ in chlorine-rich groundwater (Fig. 8c). Neoformation of the Mg-bearing clay minerals and dissolution of the 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. 8d) mimics the diamond-shaped field of the Piper diagram (Fig. 7a). Even though concentrations of Cl^- and SO_4^{2-} are equal to $HCO_3^- + CO_3^{2-}$ in the mixed groundwater, the bicarbonate-rich groundwater is enriched in bicarbonates. In accordance with seawater and thermal fluids, the chlorine-rich groundwater contains higher chloride and sulfate contents in comparison with bicarbonates. Further, the ratio of Ca^{2+} and Mg^{2+} versus $Na^+ + K^+$ represents the following decreasing order (Fig. 8d): $C1_{BGW} > C2_{MGW} > C3_{ClGW}$.

In the ilr-ion plot, Fig. 8a provides interesting information about the hydrothermal system. Most of the bicarbonate-rich groundwater samples are composed of equal chloride and sulfate contents. Their mixture with hydrothermal fluid and/or seawater results in increasing chloride, but steam-heating leads to elevated sulfate contents. Calcium concentration is higher than magnesium level in the majority of samples (Fig. 8a).

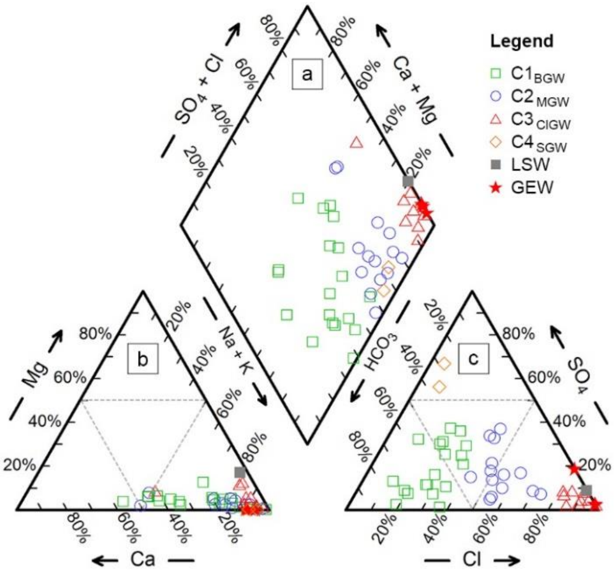


Fig. 7 Piper diagram showing composition of Phlegrean Fields groundwater, grouped according to HCA results, together with geothermal water (GEW) and local seawater (LSW): (a) the diamond-shaped field; (b) the cation field; and (c) the anion field. The abbreviations in the legend are the same with Fig. 4.

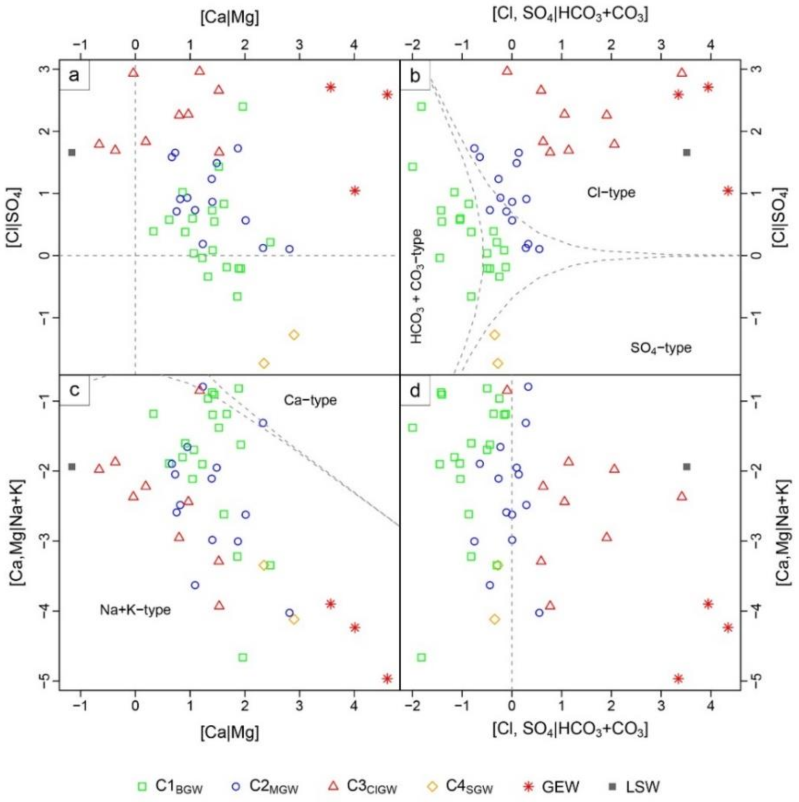


Fig. 8 Chemical composition of the collected groundwater samples on the ilr-ion plot: (a) the information in this panel is not specifically presented in the Piper diagram; (b) corresponds to the anion field of the Piper diagram; (c) corresponds to the cation field of the Piper diagram; and (d) corresponds to the diamond-shaped field of the Piper diagram. The abbreviations are the same with Fig. 4.

4.3.2. The relationship between selected variables

In order to present the importance of considering all relevant information in compositional data, strength of the relationship between Na^+–Cl^-, B–Cl^- and B–Na^+ is evaluated via symmetric coordinates and log-transformed values (Fig. 9). In the groundwater samples 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. 9 because the weakest and the strongest relationships exist between the pairs of variables in the bicarbonate- and chlorine-rich groundwater, respectively. The mixed groundwater represents a moderate R-squared value. It is noteworthy that the R^2 values are higher in Fig. 9b, d and f due to spurious correlation. For instance, approximately 80% ($p < 0.05$) of sodium variation is explained by chloride in the groundwater samples mainly derived from meteoric water ($C1_{BGW}$). It reduces to 45% ($p < 0.05$) when symmetric coordinates of Na^+ and Cl^- are utilized for the linear regression model which seems more reliable regarding the groundwater origin (Fig. 9a and b). Contrary to Fig. 9f, the B–Cl^- relationship in the chlorine-rich groundwater is not statistically significant in Fig. 9e which 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 concentrations in the Cl^--bearing high temperature springs probably explain 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. 10). The chlorine-rich groundwater can consequently be considered as the deep reservoir brine partially diluted with meteoric water. Depending on hydrogeology of the volcanic aquifer, the Na–Cl brines may be noticeably diluted with poor Li and As meteoric waters and give rise to the mixed groundwater. Seawater contribution is likely a process explaining the deviation from the mixing trend between the bicarbonate- and chlorine-rich groundwater (Fig. 10). Arsenic enrichment around Solfatara crater (the sulfate-rich groundwater spatially occurs in this area) is indicative of significant As in the vapor phase, condensing at shallow depth and giving rise to the steam-heated groundwater. 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. 11). The bicarbonate- and sulfate-rich groundwater occurs in quadrant 3 of the graph ($[Na^+|Cl^-] > 0$ and $[Na^+, Cl^-|HCO_3^-] < 0$), but the chlorine-rich groundwater lies in quadrant 1 ($[Na^+|Cl^-] < 0$ and $[Na^+, Cl^-|HCO_3^-] > 0$). Nevertheless, the mixed groundwater occurs 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. 4; Table 3) decrease as moving from quadrant 3 to quadrant 1 (Fig. 11). It verifies that the bicarbonate-rich meteoric water flows towards the sea, interacts with the chlorine-rich thermal/seawater endmember and gives rise to the mixed groundwater. The chlorine-rich groundwater is 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).

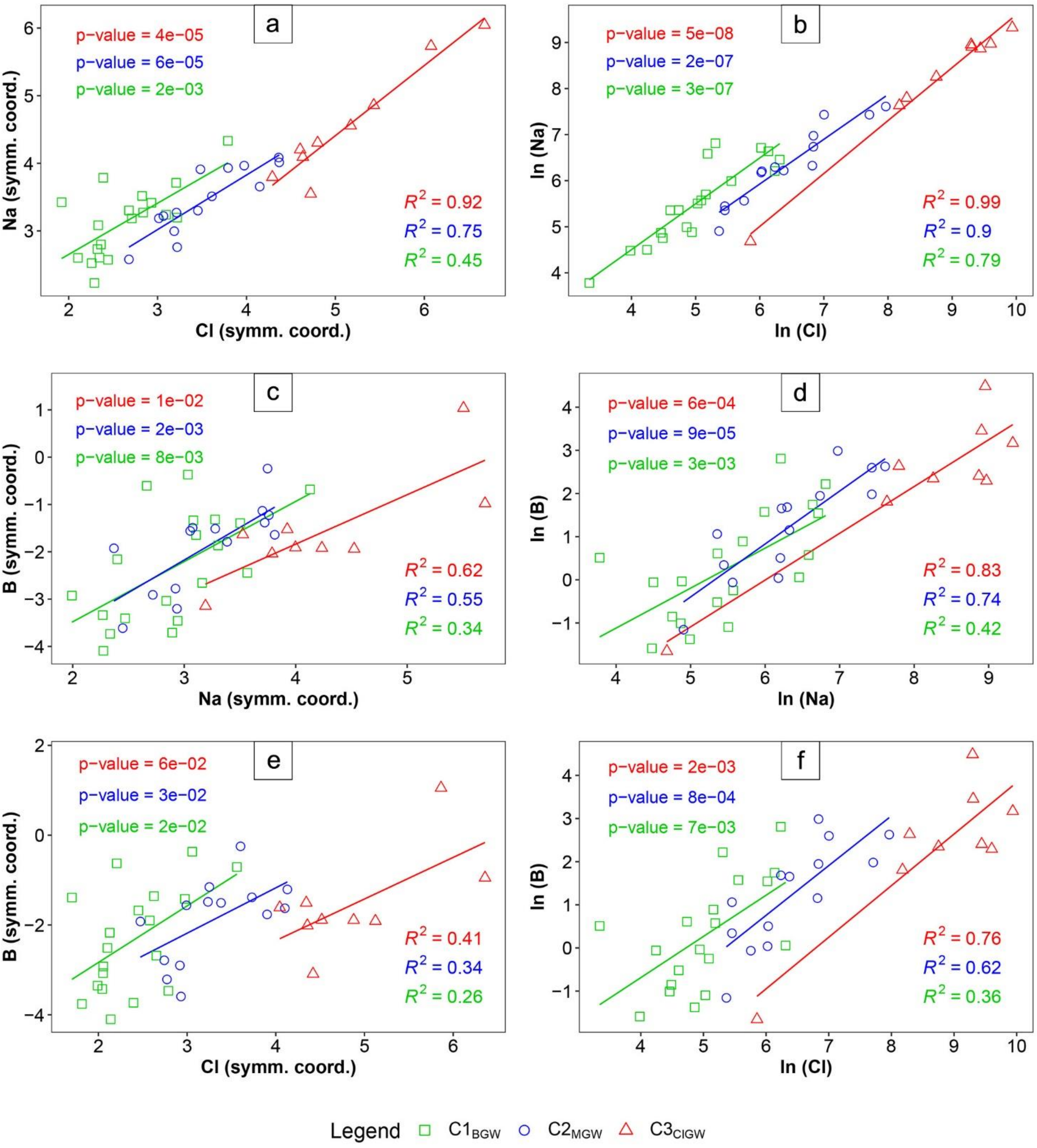


Fig. 9 The relationship between Na^+, Cl^- and B considering the symmetric coordinates (a, c and e) and the classical log-transformed data (b, d and f). The ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-}, Cl^-, NO_3^- and F^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity were used for calculation of the symmetric coordinates. The sulfate-rich groundwater is excluded because of few samples. The abbreviations in the legend are the same with Fig. 4.

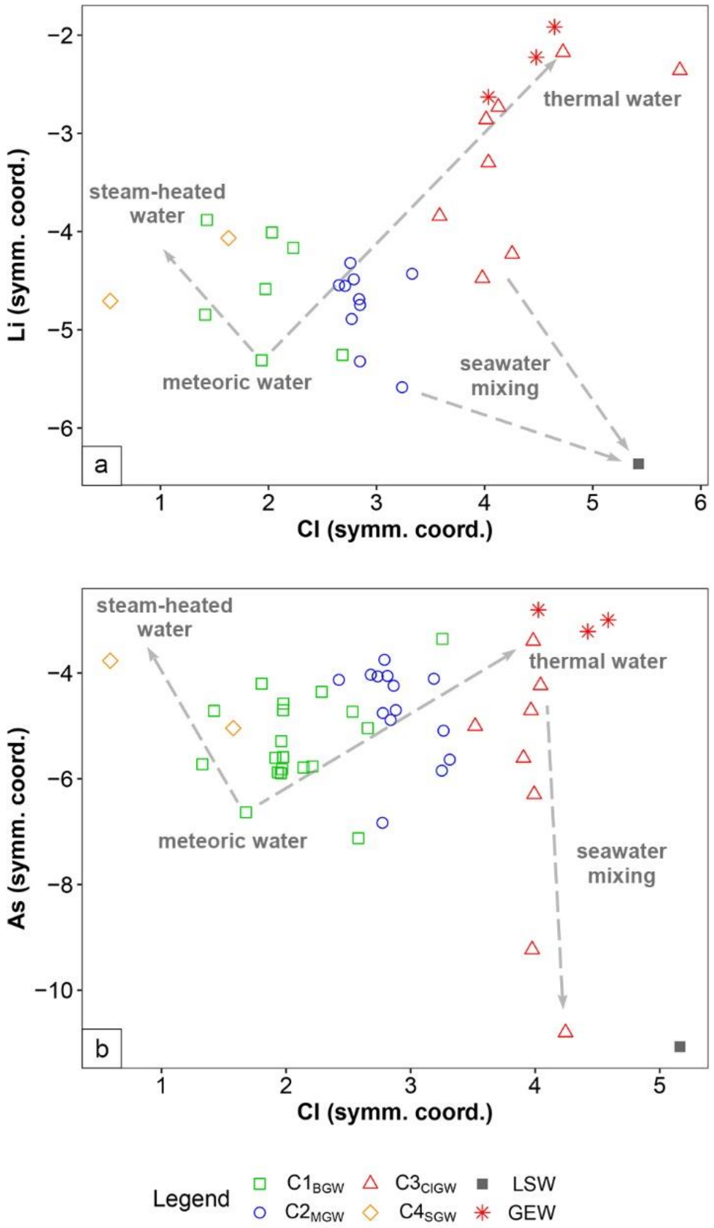


Fig. 10 Li–Cl^- and As–Cl^- scatterplots (a and b, respectively) generated by the symmetric coordinates. The ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-} and Cl^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity are used for calculation of the symmetric coordinates. The samples with a lithium content of below the detection limit are excluded in the Li–Cl^- panel. The abbreviations in the legend are the same with Fig. 4.

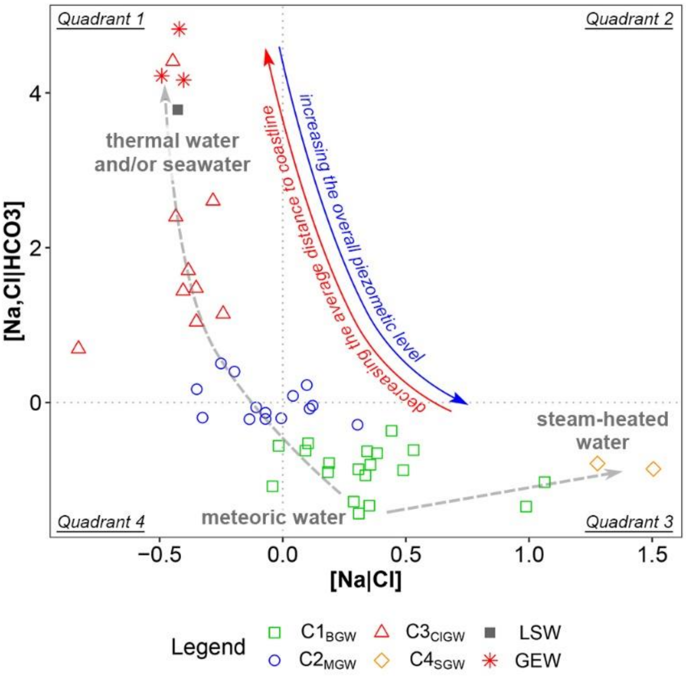


Fig. 11 The relationship between the isometric log-ratios computed using Na^+, Cl^- and HCO_3^- in mg/L. The abbreviations in the legend are the same with Fig. 4.

Precursors of volcanic eruption are mainly identified via geochemical measurements regularly on the proximal emissions such as fumaroles, hot springs and volcanic lakes (Aiuppa et al., 2007; De Moor et al., 2016) or by geodetic and seismic monitoring (Patanè et al., 2003; Sparks, 2003; McNutt and Roman, 2015). However, the intermediate or remote hydrogeochemical monitoring of volcanoes are sometimes documented for identification the precursors (Armienta and De la Cruz-Reyna, 1995; Caracausi et al., 2003). Barbieri et al. (2021) recently indicated that the remote approach is likely easier and more reliable than sampling the proximal volcanic fluid emissions. They found an increase of B and Li concentrations before the 2014 Bárðarbunga eruption (~115 km from the monitoring station) and the 2018 Mw 5.0 earthquake (occurred ~80 km from the monitoring site) in central Iceland during the 2010-2018 groundwater monitoring. It indicates the ongoing investigation on finding better precursors of volcanic eruption through hydrogeochemical monitoring. Given the results of compositional data analysis in Figs. 8 to 11, the recent advances in statistical analysis deserve further attention to explore the full potential of groundwater hydrogeochemistry for delineating volcanic activity and the related seismic events. The new approaches might identify the most useful chemical variables that can validate the seismic and volcanic events over time, avoiding arbitrary selection of a hydrogeochemical precursor. This is of the utmost importance in active volcanic areas like Phlegrean Fields, where densely populated areas are located inside and around craters.

5. Conclusions

Phlegrean Fields are an active volcanic area that has been investigated for several decades to understand the history of volcanic activity and monitor the 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 section, 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 abovementioned 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 groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The relationships between variables in each group is then explored using compositional data analysis and 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 indicates that the bicarbonate-rich meteoric water flows towards the sea, interacts with the chlorine-rich thermal/seawater endmember and gives rise to the mixed groundwater. The rising hydrothermal vapor, however, increases sulfate and arsenic contents of the groundwater body around Solfatara crater. This investigation revealed some advantages of compositional data analysis, but applying this technique is recommended in future to highlight its full potential for mitigating volcanic risk more effectively.

Contribution of the Ph.D. candidate

Collecting samples; Instrumental analysis; Determining research objectives; Data visualization; Data interpretation; Writing the original draft; Revising the draft

Section 2.2:
The Dissolved Gases in Groundwater: A Follow-up Study

Abstract

Phlegrean Fields are one of the most active volcanic areas in the world, where the spatial distribution of radon and carbon dioxide in groundwater and the statistical relationships between the dissolved gases and other variables deserve further attention to assess the potential proxies of volcanic-related phenomena. Compositional data analysis (CoDA) was proposed at the end of last century and further developed in the last decades for reliable data mining, but its potential has not been fully explored for characterization of the groundwater aquifers affected by hydrothermal activity. Based on a prospecting campaign mainly aimed at the determination of both radon and carbon dioxide in Phlegrean Fields groundwater, this section explores the spatial patterns of these gases in the local aquifer system and uses a CoDA approach to extract the relevant information and determine the meaningful geochemical associations. The results show that both dissolved gases share almost similar spatial distributions which correspond to the hydrothermal system, but the log-ratio transformed CO_2 distinguishes bicarbonate-rich groundwater better than the raw values. Principal component analysis reveals two associations: (1) Ca^{2+}, Mg^{2+}, K^{+}, SO_4^{2-}, HCO_3^{-}+CO_2 and pH; and (2) Na^{+}, Cl^{-}, As, B, Li, Rn, TDS and T. It highlights that the groundwater composition is generally influenced by two main factors: (1) meteoric water, which is modified by the CO_2-rich magmatic gas in some cases; and (2) hydrothermal fluid and/or seawater. The results are in agreement with literature and application of CoDA is recommended in future investigations because the study area is highly populated and considering compositional nature of geochemical data might help to propose new tools for characterizing volcanic eruption more efficiently.

1. Introduction

Radon is a chemically inert gas with variable levels in the geological environment, depending on concentrations of the parent isotopes in the earth's crust. Its abundance is generally higher in igneous rocks and it could be detected in soil gas and groundwater. This noble gas is colorless, tasteless and odorless, which makes it undetectable by the human senses even at high concentrations (Ciotoli et al., 2017). Radon has three natural alpha-emitting isotopes (Duggal et al., 2017 and references therein; Groves-Kirkby et al., 2006): (1) ^{222}Rn (half-life = 3.8 days) is a direct product of ^{226}Ra in the ^{238}U decay-series; (2) ^{220}Rn (half-life = 55.6 seconds) is a decay product of ^{232}Th; and (3) ^{219}Rn (half-life = 3.6 seconds) is a decay product of ^{235}U. Of the three isotopes, ^{222}Rn is the most abundant and it is customarily called "radon". However, ^{219}Rn and ^{220}Rn are usually referred to as actinon and thoron, respectively. This section deals only with ^{222}Rn, which will be referred to as radon (Rn) from now on.

Chirkov (1975) proposed anomalous radon levels as the precursor of a volcanic eruption. In active volcanic areas, variations in Rn concentration could be a result of magma ascent, changes in temperature and/or depth of the hydrothermal systems, stress and strain degree in rocks, opening or sealing of cracks/fissures of gas conduits, and increased thermo-mechanical waves (İçhedef et al., 2020; Neri et al., 2016; Segovia, 1991). Significant Rn release was recorded after rocks underwent the processes related to magma injection such as heating, deforming and fracturing in the laboratory (Koike et al., 2015; Sakoda et al., 2011 and references therein). Soil gas compositions showed that Rn concentration peaks during the eruption are at least four times the background values in four explosive American stratovolcanoes (Segovia and Mena, 1999). However, radon degassing could completely occur when major gas species (i.e. CO_2 and H_2O) act as radon carriers (Gauthier et al., 1999). It highlights that a large amount of CO_2 (the most abundant species in a water-free magmatic gas) is similarly released into the atmosphere during eruptions and quiescent periods (Granieri et al., 2014). The CO_2 directly emitted from active volcanoes is mostly originated from the CO_2 dissolved in the mantle, the recycled CO_2 from subducted crustal material (Marty and Tolstikhin, 1998) and decarbonation of the shallow crustal material (Troll et al., 2012). Isotopic composition of carbon could enlighten the relative contribution of mantle and crust (Chiodini et al., 2011) and demonstrate the behavior of volcanoes over time because magmatic intrusions may interact with crustal material during eruption and noticeably increase the CO_2 output of the volcanic system (Troll et al., 2012). The CO_2 degassing in quiescent volcanos occurs through fumaroles, hot springs and volcanic lakes around craters and micro-seepages from their flanks. Following carbon dioxide dissolution in water, pH-dependent reactions maintain the equilibrium between carbonate species. Several studies observed a close relationship between faults and the elevated concentrations of Rn and CO_2 which demonstrates that the generally high permeable zones around crustal fractures and discontinuities favor the upward migration of gas from the deep crust (Segovia, 1991; Kumar et al., 2016; Terray et al., 2020).

In southern Italy, previous studies in soil gas and groundwater revealed that Rn concentrations in Phlegrean Fields and Ischia (>10 Bq/L on average) are higher than those of Vesuvius and Vulcano (<10 Bq/L on average; Avino et al., 1999). A recent regional-scale investigation also confirmed association between the greatest Rn flux from soil and the volcanic centers, including Phlegrean Fields (Guarino et al., 2022). Due to the occurrence of bradyseismic crises in Phlegrean Fields at the end of the twentieth century, groundwater monitoring was performed for investigating spatial and temporal variations of the dissolved Rn levels between 1983 and 1997 (Capaldi et al., 1992; Tedesco et al., 1996; Avino et al., 1999). The results of an eighteen-year soil CO_2 flux survey at Solfatara crater, the present epicenter of volcanic unrest, showed that the area releasing deep-sourced CO_2 was tripled between 2003 and 2016, being consistent with the increased injection of magmatic fluids into the hydrothermal system. The elevated background CO_2 flux indicated that magmatic carbon dioxide widely affects both Solfatara crater and the surrounding environment (Cardellini et al., 2017). Although degassing and seismic activity were concentrated around the Solfatara area (Fig. 1) before 1538 CE, they increased near Monte Nuovo several days prior to the eruption. This change in volcanic-related phenomena also shows the importance of extending the existing monitoring plans to areas other than the Solfatara crater (Di Napoli et al., 2016 and references therein). Further, when relative ratios (i.e. the majority of geochemical data) rather than absolute values are of interest, data analysis in the classical Euclidean space will deliver faulty results. These data are referred to as compositional data and considering their own Euclidean geometry on the simplex is critical (Aitchison, 1982, 1986). To the best of our knowledge, this aspect has not gained enough attention for studying groundwater hydrogeochemistry in the aquifers affected by hydrothermal activity and most recent surveys determined radon and CO_2 variations in soil gas and fumaroles of specific monitoring points (e.g., Aiuppa et al., 2013; Cardellini et al., 2017; Chiodini et al., 2021; Pedone et al., 2014; Sabbarese et al., 2020). Hence, the objectives of this study are to capture the spatial distribution of Rn and CO_2 in the central part of the Phlegrean Fields caldera together with respect compositional data structure for investigating the statistical relationship between the dissolved gases, major ions, Li, As, B and other physicochemical characteristics of the groundwater body.

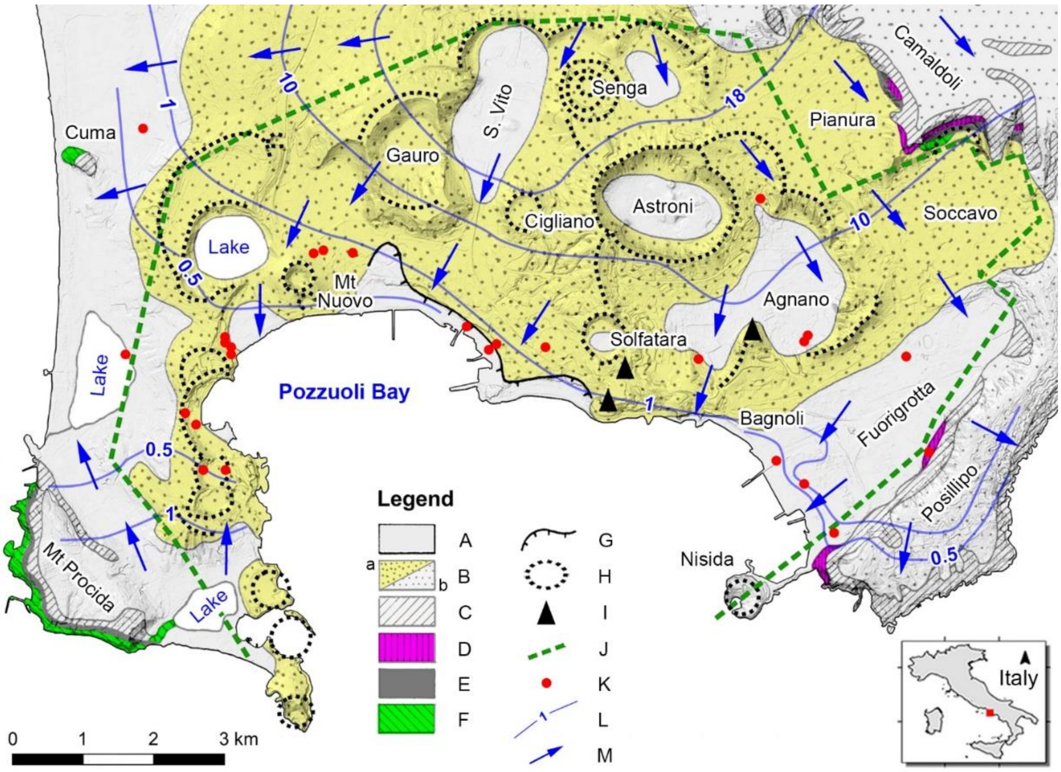


Fig. 1 Map of Phlegrean Fields with location of the sampling points (modified after Orsi et al., 1996): (A) Recent active continental sediments; (B) Volcanics younger than 12 ka: "a" and "b" show proximal and distal deposits, respectively; (C) Neapolitan Yellow Tuff (NYT; 12 ka B.P.; (D) Volcanics erupted 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) Groundwater sampling points; and (L) Piezometric contour lines (m a.s.l.; De Vita et al., 2018); and (M) Groundwater flow direction.

2. Study area

The Phlegrean Fields, south Italy, refer to a large active volcanic system (about 12 km in diameter; Fig. 1) along the coastline of Tyrrhenian Sea. Campanian Ignimbrite and Neapolitan Yellow Tuff explosive eruptions (40 and 15 ka B.P., respectively) influenced the present morphology (Natale et al., 2021 and references therein). There were several centuries of deflation after the last eruption at Monte Nuovo (Fig. 1) in 1538 CE, but land deformation and seismic unrest since the 1950s led to bradyseismic crises at the end of the last century and highlighted the importance of volcanological studies to prepare for an emergency. The present phase of ground uplift started from 2004–2005 which explains the frequent shallow (<2 km) earthquakes together with significant variation in degassing rate and chemical composition of the fumaroles. Therefore, the resurgent sector of Neapolitan Yellow Tuff (NYT) caldera has been extensively investigated in the last decades. The high magmatic CO_2 degassing emissions (ca. 1600 ton/day) in soils (Chiodini et al., 2001, 2012) and fumaroles (Aiuppa et al., 2013, 2015; Pedone et al., 2014) of the Solfatara-Pisciarelli hydrothermal area were suggested as a significant factor in the uplift and seismicity (De Siena et al., 2010; D'Auria et al., 2012; Chiodini et al., 2015). The recently measured total CO_2 flux of up to 3000 ton/day at Solfatara volcano is comparable with the average volcanic plume fluxes of persistently degassing active volcanoes (Cardellini et al., 2017).

Drillings for geothermal exploration up to about 3 km depth indicated that the hydrogeological system of Phlegrean Fields is quite complicated. The permeable horizons occur a few hundred meters below the surface, except for the fractured areas. The locally superimposed aquifers can be considered as a unique aquifer at the basin scale because of lateral and vertical discontinuities of low-permeability horizons and geometrically variable lithology. The groundwater body is mainly recharged by meteoric water, but the mixing of meteoric water, seawater and deep geothermal fluids occur to different extents (Capaldi et al., 1992 and references therein). Around Solfatara volcano, the CO_2-rich magmatic gas that flashes and vaporizes the hydrothermal liquids of meteoric origin, with fixed CO_2 fugacity due to fluid-rock interactions, generates a gas plume in a deep zone with temperatures around 360 °C. Then, re-equilibration of H_2 and CO occurs in the rising pure vapor phase at about 200 to 240 °C (Caliro et al., 2007). From the hydrogeological point of view, the radial groundwater flow is oriented towards the coastline in the southern and western sectors, and locally feeds the coastal lakes and mud pools (Fig. 1).

3. Materials and methods

3.1. Sampling and field measurements

Because the highest Rn values (>40 Bq/L) in groundwater were observed in the central part of the Phlegrean Fields caldera (Avino et al., 1999), twenty-six groundwater samples from thermal springs and private wells that mostly occur inside NYT caldera and a seawater sample from Porto di Baia (Pozzuoli Bay) were collected in November and December 2020. Well purging was performed before sampling to ensure that the samples are representative of the groundwater in the vicinity of the well. At each station, two aliquots were collected following 0.45-μm filtration and HNO_3 acidification for determination of cations and metal(loid)s, but four unfiltered and unacidified aliquots were collected for anions, alkalinity, radon and CO_2 measurements. The sampling bottles were rinsed three times with the water from the exact site of sampling prior to filling them completely which leaves no headspace and prevents the dissolved gases from escaping into the air. Airtight sealed containers were utilized for the aliquots that underwent Rn and CO_2 determinations. All containers were made of high-density polyethylene, but the aliquot for Rn measurement was stored in a glass bottle. Total dissolved solids (TDS), pH and temperature were recorded in situ with portable devices.

3.2. Laboratory analysis

On the same day of sample collection, alkalinity, Rn and CO_2 were determined at the Environmental Geochemistry Laboratory of the Department of Earth, Environmental and Resources Sciences (DiSTAR), University of Naples Federico II (UNINA). Alkalinity was measured by titration of 100 ml of a given sample with HCl (0.1 N) and the values are expressed as HCO_3^-. Radon concentrations were measured with the AquaKIT accessory (which is composed of a degassing vessel, a security vessel, a pump and tubes) for the radon monitoring system AlphaGUARD PQ2000 PRO. Briefly, the “empty” measurement set-up is assembled, AlphaPUMP is turned on with the flow rate of 1 L/min and the system is rinsed with room air for 10 min. The gas cycle is then closed, AlphaPUMP is turned on with the flow rate of 0.5 L/min, AlphaGUARD is switched on in “FLOW, 1 min” operation mode and the background Rn level is detected in the “empty” measurement set-up for 10 min. Subsequently, the sample is slowly injected into the system, the gas cycle is closed immediately to prevent the uncontrolled escape of radon out of the measuring set-up, AlphaPUMP is turned on with the flow rate of 0.5 L/min, AlphaGUARD is switched on in “FLOW, 1 min” operation mode and the dissolved Rn level is detected for 20 min. Finally, water temperature is determined and Rn concentration (c_{Water}; Bq/L) in groundwater sample is calculated as follows:

$$c_{Water} = \frac{c_{Air} \times \left(\frac{V_{System} - V_{Sample}}{V_{Sample}} + k\right) - c_0 \times \left(\frac{V_{System} - V_{Sample}}{V_{Sample}}\right)}{1000} \tag{1}$$

where c_{Air} is Rn concentration (Bq/m³) in the measurement set-up after sample introduction and expelling the dissolved gas, c_0 is the background Rn concentration (Bq/m³) in the measurement set-up before sample introduction, V_{System} is the interior volume (ml) of the measurement set-up, V_{Sample} is the volume (ml) of the introduced water sample, and k is Rn distribution coefficient water/air obtained from the temperature dependency curve of Rn diffusion coefficient based on water temperature (with falling temperature, the amount of soluble Rn in water and the k-value increase). In this study, V_{System} and V_{Sample} are 1102 and 100 ml, respectively. It is worth mentioning that AlphaGUARD (the lower measuring range is 2 Bq/m³) records a value every minute and the time period of almost stable readings is considered to compute the average c_{Air} or c_0 value for a given sample. The instrument was calibrated with reference to national radon standards (NIST, NPL and PTB) and the analytical precision of 11% was obtained by replicate measurements.

Considering a maximum of 6-hour gap between sampling and measurement, the values obtained from Eq. (1) may be corrected based on the percentage of disintegrated atoms in this time span (Landauer, 1961):

$$N = N_0\, e^{-\lambda T} \tag{2}$$

where N_0 is the number of atoms at time zero, N is the number of atoms remaining after time T and λ is the disintegration constant per hour computed via Eq. (3):

$$\lambda = \frac{0.693}{\mathrm{T}_{1/2}} = \frac{0.693}{3.825 \text{ days} \times 24 \text{ h}} = 0.007549 \text{ per hour} \qquad (3)$$

where $\mathrm{T}_{1/2}$ is radon half-life (3.825 days). In Eq. (2), if N_0 is taken as 100, λ as 0.007549 per hour and T as 6 hours, then N equals 95.6, and 4.4 atoms out of 100 (or 4.4%) have decayed until laboratory analysis. Thus, $1.046 \times c_{\mathrm{Water}}$ accounts for the representative dissolved Rn quantity in groundwater. The measurements indirectly show ^{226}Ra disintegration at depth and in the groundwater aquifer.

Formula (1) is also applicable to the dissolved CO_2 content, but carbon dioxide concentration needs to be detected via ECOPROBE 5 and some parameters in the abovementioned formula should be adjusted. Briefly, after rinsing the "empty" measurement set-up with room air using AlphaPUMP with the flow rate of 1 L/min for 10 min, the gas cycle is closed, ECOPROBE 5 is turned on with the flow rate of 0.5 L/min and the background CO_2 level is detected in the "empty" measurement set-up for 3 min. Subsequently, the sample is slowly injected into the system, the gas cycle is immediately closed to minimize uncontrolled escape of the dissolved gas out of the measuring set-up, ECOPROBE 5 is turned on with the flow rate of 0.5 L/min and the dissolved CO_2 level is detected for 3 min. Finally, water temperature is determined and CO_2 concentration (c_{Water}; mg/L) in water sample is calculated via formula (1). For carbon dioxide calculation via the formula, c_{Air} is CO_2 concentration (mg/m³) in the measurement set-up after sample introduction and expelling the dissolved gas, c_0 is the background CO_2 concentration (mg/m³) in the measurement set-up before sample introduction, V_{System} is the interior volume (ml) of the measurement set-up, V_{Sample} is volume (ml) of the introduced water sample, and k is CO_2 distribution coefficient water/air obtained from the temperature dependency curve of CO_2 diffusion coefficient (Carroll and Mather, 1991) based on water temperature. In this study, V_{System} and V_{Sample} are 404 and 100 ml, respectively. It is noteworthy that ECOPROBE 5 (the detection limit is 36 mg/m³) records a value every 0.5 seconds. A single-point calibration, controlled by the instrument's internal program, was implemented using the LINDE calibration standard cylinder that contains 100 ppm Isobuthylene mixed into synthetic air. Repeated measurements showed an analytical precision of 5% in the present study.

Major ions and Li were determined via ion chromatography in the laboratory of fluid geochemistry at the INGV-Osservatorio Vesuviano, Napoli, Italy. However, As and B were measured with inductively coupled plasma-optical emission spectroscopy (ICP-OES) at DiSTAR (UNINA). Blank, duplicate, and certified standard solutions were analyzed to check quality of the analytical results. Measurement accuracy is better than 5%. The relative standard deviation of analytical duplicates is also below 5%, except for As and B (23 and 24%, respectively).

3.3. Statistical analysis

Estimating the precision of every statistic helps to obtain confidence intervals and test hypothesis, but the traditional methods tend to overestimate width of the confidence interval. Bootstrapping is a mathematically simple alternative that makes fewer assumptions and can find the confidence interval for any statistic more accurately (Wright et al., 2011 and references therein). In this study, based on 10000 bootstrap replications, significance of the difference in average radon values between chlorine- and bicarbonate-rich groups is investigated. The R packages "simpleboot" (Peng, 2019) and "boot" (Canty and Ripley, 2021) are applied for bootstrapping and assessing significance of the difference.

Hydrogeochemical data are mainly non-negative values which are compositional by nature. It means that each component accounts for a proportion of some whole. The Euclidean geometry on the simplex needs to be considered instead of the classical Euclidean geometry in real space for statistical analysis of compositional data (see Aitchison, 1986; Ebrahimi et al., 2022), otherwise faulty results arise. The notion of spurious correlation between the components was expressed at the end of the nineteenth century (Pearson, 1897) and it was underlined in a geological context later (e.g., Chayes, 1960). Comprehensive examination of this issue led to the introduction of additive log-ratio- (alr-), centered log-ratio- (clr-) and isometric log-ratio- (ilr-) transformations (Aitchison, 1982, 1986; Egozcue et al., 2003) for dealing with the data problem. These techniques are termed compositional data analysis altogether. In the present study, descriptive statistics are provided for raw data and clr-transformed values because dealing with the constant sum constraint of compositional data is of the utmost importance in univariate data analysis (Reimann et al., 2012). The clr-transformation is performed through dividing each variable by the geometric mean of the whole set of variables, followed by a natural log transformation (Aitchison, 1982, 1986). In addition, principal component analysis (PCA) is performed after clr transformation of the compositional parts to summarize the main information in the multivariate data structure. It reduces dimensionality by constructing the new coordinates (i.e. principal components) that express the highest variability in the dataset. Results of PCA are finally represented via a biplot, a two-dimensional visualization of observations and variables, along with a map that shows the spatial distribution of PC1 scores. Further details about performing PCA in the context of compositional data analysis together with the biplot interpretation are available in Filzmoser et al. (2018) and Somma et al. (2021). The PCA is performed with the R package “robCompositions” (Filzmoser et al., 2018).

4. Results and discussion

4.1. Composition of the groundwater body

Results of the laboratory and field measurements are presented in Table 1. To understand statistical distribution measures for each variable in the whole dataset, the percentiles of raw data are also provided. It is worth mentioning that the statistics (such as average and standard deviation) that are based on the Euclidean distance are excluded. However, as discussed in the previous section, compositional data (i.e. the ions, Li, As, B and CO_2) should be considered in the Euclidean structure of the simplex even for univariate analysis. Contrary to the logarithmic transformation, a log-ratio transformation can change percentiles of a distribution (Reimann et al., 2012) which likely makes the descriptive statistics invalid for the compositional data in Table 1. Thus, percentiles of the clr-transformed data are also reported in Table 2. A clr-transformed value in Table 2 explains all relative information of a given variable to the remaining variables. In other words, it indicates dominance of a variable in a composition. The transformation generates dimensionless numbers for the dissolved species in groundwater which is not suitable for comparison with literature values, but median and variance of the single clr-transformed variables are still comparable. In fact, the information in Tables 1 and 2 are not much different considering that one is looking at different and dimensionless numbers in Table 2. Order of the dissolved species does not change after transformation and the species with a higher concentration is characterized by a greater clr-transformed value (Reimann et al., 2012). The median values follow the decreasing order of $\mathrm{clr}(\mathrm{HCO_3^-}) > \mathrm{clr}(\mathrm{Cl^-}) > \mathrm{clr}(\mathrm{SO_4^{2-}})$ for anions and of $\mathrm{clr}(\mathrm{Na^+}) > \mathrm{clr}(\mathrm{Ca^{2+}}) > \mathrm{clr}(\mathrm{K^+}) > \mathrm{clr}(\mathrm{Mg^{2+}})$ for cations (Table 2), being consistent with Table 1. The median $\mathrm{clr}(\mathrm{HCO_3^-})$ of 3.4 is the greatest value which is about $2.7^{3.4} = 29.3$ times the geometric mean of all components and the median $\mathrm{clr}(\mathrm{Na^+})$ is at least twice the value for the other cations. Of the compositional variables, only the median values for magnesium, lithium, arsenic and boron are below the geometric mean of all parts. The interquartile range (IQR) and median absolute deviation (MAD) are estimations for variance of the clr-transformed data. The variable with higher variance has more influence on the multivariate data ensemble (Reimann et al., 2012). In Phlegrean Fields, relative concentrations of chloride, carbon dioxide, magnesium, boron, calcium and arsenic represent the highest variation (see IQR and MAD in Table 2). Therefore, $\mathrm{clr}(\mathrm{Cl}^-)$, $\mathrm{clr}(\mathrm{Mg}^{2+})$, $\mathrm{clr}(\mathrm{Ca}^{2+})$, clr($CO_2$), clr(As) and clr(B) have more influence than $\mathrm{clr}(\mathrm{HCO}_3^-)$ and $\mathrm{clr}(\mathrm{Na}^+)$, the dominant compositional parts considering the P50 values (Table 2).

The prevalent ion species has been regularly identified based on concentrations in meq/L, but Table 3 shows that the clr-transformation of ions (in meq/L) could also be used for the same purpose in the context of compositional data analysis. The samples are categorized in bicarbonate-, chlorine- and sulfate-rich groups (BG, ClG and SG, respectively) on the basis of the dominant anion species (Fig. 2, Tables 1 and 3). The groundwater is generally enriched in bicarbonate species towards east, but chloride concentration exceeds that of HCO_3^- towards west (Fig. 2). Although raw values show that concentrations of bicarbonate and calcium are almost similar in both groups, clr-transformation clarifies that the bicarbonate-rich group contains relatively higher levels of the ions (Fig. 3). In accordance with Table 2, $\mathrm{clr}(\mathrm{Mg}^{2+})$, clr(Li), clr(As) and clr(B) are below and rest of the compositional parts are above the geometric mean of the composition in both groups. The $\mathrm{clr}(\mathrm{Cl}^-)$, $\mathrm{clr}(\mathrm{Na}^+)$, clr(Li), clr(As), clr(B), Rn, temperature and TDS are generally greater in the chlorine-rich group, highlighting higher contribution of the deep hydrothermal system and/or seawater (Fig. 3). The difference in average Rn concentration between the bicarbonate- and chlorine-rich groups (Fig. 4) is statistically significant because the middle 95% of the distribution is different from zero. Radon concentration is the lowest in the local seawater and the highest in Stufe di Nerone (<D.L. and >220 Bq/L, respectively; Table 1) in the western sector of Phlegrean Fields (Fig. 2). The groundwater temperature of above 55 °C may also confirm existing of the pathways for upward migration of the hydrothermal fluids in this site. However, in accordance with Capaldi et al. (1992), the spatially variable Rn concentrations in the study area corresponds to the site-specific characteristics of the groundwater aquifer. Avino et al. (1999) indicated seasonal dependency of Rn measurements based on groundwater monitoring over 15 years. They reported that radon content in groundwater varies from 0.9 to 321.5 Bq/L, being consistent with results of the present study (1.1-465.2 Bq/L). The recorded pH values range from 6.16 to 9.14 (Table 1) and the bicarbonate-rich group mostly contains the groundwater samples with pH > 7 (Fig. 3).

Table 1 Coordinates (UTM zone 33N) of the collected groundwater samples (N=26) together with their chemical composition and physicochemical properties. The samples are collected from wells, unless otherwise specified. The descriptive statistics are obtained after zero replacement. The local seawater (W27) is excluded for calculating the descriptive statistics. Min: minimum; P: percentiles (P50=median); Max: maximum; MAD: median absolute deviation; and MAD.log: median absolute deviation for the log-transformed values.

	Sample ID	**Descriptive ID**	**X**	**Y**	HCO_3^-	Cl^-	SO_4^{2-}	Na^+	K^+	Mg^{2+}	Ca^{2+}	CO_2 [a]
Bicarbonate-rich group (BG)	W1	Carannante	428750	4519699	1101.8	504.1	596.0	514.1	131.0	39.5	306.8	624.4
	W2	Castello di Baia	422096	4518143	887.8	417.7	396.5	545.8	68.7	78.8	104.2	72.8
	W3	Conte	423892	4521197	687.2	121.0	69.1	235.8	85.6	2.5	54.5	50.6
	W4	CUS	431990	4518391	294.5	76.6	162.5	109.6	45.1	16.7	89.6	12.0
	W5	Di Dio	429621	4521958	690.2	86.2	165.1	123.7	68.6	12.0	180.2	69.0
	W6	Hotel Neronensis	425916	4519903	426.2	55.5	10.0	22.1	17.0	17.7	111.2	28.8
	W7	INVITALIA–4	429838	4518276	425.6	78.9	113.7	167.3	33.6	12.6	43.9	56.8
	W8	Kyme marmi	420944	4522943	380.5	142.6	82.9	212.0	36.2	3.4	37.4	19.2
	W9	M. Oltremare	431667	4519742	349.4	98.9	136.8	135.6	58.2	12.2	70.5	16.6
	W10	Serapeo (spring)	425807	4519836	1071.9	408.2	239.0	566.2	122.9	37.2	100.2	167.0
	W11	Sud Cantieri	425483	4520157	378.0	103.0	134.6	166.4	56.3	14.1	58.0	8.0
	W12	Tortorelli	426602	4519870	1813.4	167.1	585.1	927.9	207.3	4.5	29.7	202.4
Chlorine-rich group (ClG)	W13	Agriturismo fondi di Baia	421793	4518136	243.3	210.6	193.6	135.4	51.0	25.1	132.6	n.m.
	W14	Centro Ippico	423344	4521189	504.9	336.2	196.9	432.8	107.8	9.0	30.5	18.2
	W15	De Pisis	430258	4519985	1504.9	2466.4	339.2	1846.8	270.6	62.2	226.5	408.8
	W16	Esposito	423478	4521232	497.6	344.6	214.6	473.3	103.0	2.7	26.1	13.6
	W17	Franco Lo Scoglio	422175	4519840	912.2	17661.4	1915.8	10261.6	450.8	394.9	305.3	343.8
	W18	Franco Lo Scoglio (spring)	422193	4519766	843.9	18031.1	2021.6	11303.8	643.0	796.2	493.6	289.6
	W19	Grotta dell'Acqua (spring)	420700	4519762	801.8	5229.9	173.0	3430.3	129.5	<D.L.	38.6	161.8
	W20	INVITALIA–15	430239	4517944	442.1	273.2	120.4	270.2	55.9	20.2	57.7	85.0
	W21	Pozzanghera Terme di Agnano	430283	4520021	888.4	640.4	136.8	589.7	103.2	25.0	102.8	925.2
	W22	Stufe di Nerone–1	422100	4519925	312.2	11315.9	462.5	7194.8	151.9	7.0	108.2	29.0
	W23	Stufe di Nerone–2	422081	4519944	422.6	12103.8	648.3	7718.8	171.9	26.6	133.5	33.8
	W24	Tempio di Venere	421692	4518783	362.8	864.5	221.3	614.6	56.2	7.4	112.7	13.2
	W25	Terme di Baia	421538	4518942	225.0	2416.5	105.6	1511.8	93.1	10.1	119.8	8.8
Sulfate-rich group (SG)	W26	INVITALIA–30	430662	4517254	120.1	110.2	315.7	157.1	50.1	1.8	62.7	0.0
Local seawater	W27	Pozzuoli Bay (Porto di Baia)	421746	4518796	178.0	19742.7	2617.7	11509.0	382.3	1348.3	414.6	0.04
Min					120.1	55.5	10.0	22.1	17.0	0.1	26.1	0.0
P5					229.6	77.2	72.6	113.1	34.3	2.0	29.9	2.0
P25					366.6	112.9	135.2	166.6	56.0	7.1	55.3	14.3
P50					469.8	340.4	195.3	493.7	89.4	13.4	101.5	42.2
P75					876.8	2028.5	382.2	1365.8	130.6	26.2	129.4	165.7
P95					1404.1	16272.0	1598.9	9625.9	405.8	315.9	306.4	570.6
Max					1813.4	18031.1	2021.6	11303.8	643.0	796.2	493.6	925.1
MAD					324.5	367.5	126.9	515.0	57.5	15.3	64.7	47.5
MAD.log					0.3	0.7	0.3	0.7	0.3	0.4	0.4	0.8

Table 1 (continued)

	Sample ID	**Descriptive ID**	**X**	**Y**	**Li**	**As**	**B**	**Rn**	**TDS**	**pH**	**T**
Bicarbonate-rich group (BG)	W1	Carannante	428750	4519699	0.43	1.005	21.96	43.2	2010	6.16	35
	W2	Castello di Baia	422096	4518143	<D.L.	0.050	1.48	5.6	1660	7.48	21
	W3	Conte	423892	4521197	<D.L.	0.135	0.32	1.1	712	8.06	22
	W4	CUS	431990	4518391	<D.L.	0.020	0.22	1.6	567	8.05	13
	W5	Di Dio	429621	4521958	<D.L.	0.052	0.30	5.5	779	7.58	15
	W6	Hotel Neronensis	425916	4519903	<D.L.	0.029	0.11	4.2	387	7.50	18
	W7	INVITALIA–4	429838	4518276	<D.L.	0.042	0.38	18.1	435	7.08	19
	W8	Kyme marmi	420944	4522943	<D.L.	0.037	0.52	14.3	600	7.70	19
	W9	M. Oltremare	431667	4519742	<D.L.	0.032	0.28	8.6	585	7.64	15
	W10	Serapeo (spring)	425807	4519836	0.11	0.229	2.17	53.6	1570	7.00	23
	W11	Sud Cantieri	425483	4520157	<D.L.	0.137	0.48	1.8	595	7.88	16
	W12	Tortorelli	426602	4519870	0.42	0.502	9.42	15.8	2110	6.73	41
Chlorine-rich group (ClG)	W13	Agriturismo fondi di Baia	421793	4518136	<D.L.	0.011	0.24	n.m.	740	7.02	18
	W14	Centro Ippico	423344	4521189	0.15	0.600	0.92	5.8	1200	8.06	33
	W15	De Pisis	430258	4519985	1.21	0.418	14.98	11.6	5160	6.67	48
	W16	Esposito	423478	4521232	0.16	0.626	0.84	8.8	1200	8.28	34
	W17	Franco Lo Scoglio	422175	4519840	0.70	3.846	31.61	38.7	9360	6.31	52
	W18	Franco Lo Scoglio (spring)	422193	4519766	2.08	4.008	39.08	26.1	9200	6.50	51
	W19	Grotta dell'Acqua (spring)	420700	4519762	0.81	4.034	26.13	25.0	5350	6.81	31
	W20	INVITALIA–15	430239	4517944	<D.L.	0.019	0.46	20.0	855	7.06	17
	W21	Pozzanghera Terme di Agnano	430283	4520021	0.41	0.088	5.79	38.8	1760	6.24	19
	W22	Stufe di Nerone–1	422100	4519925	2.76	8.684	46.39	465.2	7950	7.25	60
	W23	Stufe di Nerone–2	422081	4519944	2.42	8.967	48.55	221.4	7380	7.37	57
	W24	Tempio di Venere	421692	4518783	0.14	0.215	2.00	9.5	1850	7.74	19
	W25	Terme di Baia	421538	4518942	0.65	0.366	5.86	10.5	3680	7.39	20
Sulfate-rich group (SG)	W26	INVITALIA–30	430662	4517254	<D.L.	0.011	0.40	20.1	589	9.14	19
Local seawater	W27	Pozzuoli Bay (Porto di Baia)	421746	4518796	<D.L.	0.026	6.45	<D.L.	>9999	8.27	16
Min					0.01	0.011	0.11	1.1	387.0	6.16	13.0
P5					0.02	0.013	0.23	1.6	468.0	6.26	15.0
P25					0.07	0.038	0.38	6.5	628.0	6.86	18.3
P50					0.13	0.176	1.20	14.3	1385.0	7.38	20.5
P75					0.60	0.620	13.59	25.8	3287.5	7.73	34.8
P95					2.34	7.522	44.56	179.4	8887.5	8.23	55.8
Max					2.76	8.967	48.55	465.2	9360.0	9.14	60.0
MAD					0.14	0.239	1.43	13.0	1167.5	0.65	7.4
MAD.log					0.75	0.948	0.96	0.5	0.5	-	0.2

All units are in mg/L, except for Rn (Bq/L), pH (unitless) and T (°C). The detection limits (D.L.) are 0.1 mg/L for the major ions and lithium, 0.003 mg/L for As and B, and 0.001 Bq/L for Rn. [a] Due to probable CO_2 loss between sampling and measurement, the detected CO_2 values are multiplied by 2 to be more consistent with the PHREEQC calculations.

Table 2 Descriptive statistics of the clr-transformed compositional data after zero replacement. For the sake of brevity, the clr-transformed values for each sample is not listed here, but utilized for generating the boxplots in Fig.3. All values are dimensionless. Min: minimum; P: percentiles (P50=median); Max: maximum; IQR: interquartile range; and MAD: median absolute deviation.

Variable	$\mathbf{clr}(\mathbf{HCO_3^-})$	$\mathbf{clr}(\mathbf{Cl^-})$	$\mathbf{clr}(\mathbf{SO_4^{2-}})$	$\mathbf{clr}(\mathbf{Na^+})$	$\mathbf{clr}(\mathbf{K^+})$	$\mathbf{clr}(\mathbf{Mg^{2+}})$	$\mathbf{clr}(\mathbf{Ca^{2+}})$	$\mathbf{clr}(\mathbf{CO_2})$	**clr(Li)**	**clr(As)**	**clr(B)**
Min	0.9	1.3	0.5	1.3	0.2	-6.5	-0.4	-15.9	-6.0	-6.7	-4.0
P5	1.0	1.8	1.1	1.9	0.3	-2.7	-0.3	-8.3	-5.9	-6.4	-3.9
P25	2.6	2.2	1.5	2.6	0.8	-1.2	0.3	-0.5	-5.7	-5.6	-3.4
P50	3.4	2.7	2.1	2.9	1.1	-0.4	1.3	0.5	-5.1	-4.9	-2.7
P75	3.7	4.1	2.4	3.6	1.5	0.3	1.5	1.5	-4.7	-4.4	-1.9
P95	4.3	4.5	3.8	4.1	2.5	1.1	3.5	1.9	-3.8	-2.7	-0.9
Max	4.3	4.6	5.2	4.5	3.4	2.0	3.7	3.0	-3.0	-2.6	-0.7
IQR	1.0	2.0	0.9	1.0	0.7	1.5	1.2	1.9	1.0	1.2	1.5
MAD	0.8	1.0	0.6	0.6	0.5	1.2	1.1	1.5	0.8	1.0	1.1

Table 3 Groundwater classification based on the dominant anion species. The prevalent anion is underlined regarding the concentrations (meq/L) of anions and their clr-transformed values (dimensionless).

	Sample ID	**Descriptive ID**	$\mathrm{clr}(\mathrm{HCO_3^-})$	$\mathrm{clr}(\mathrm{Cl^-})$	$\mathrm{clr}(\mathrm{SO_4^{2-}})$	$\mathrm{HCO_3^-}$	$\mathrm{Cl^-}$	$\mathrm{SO_4^{2-}}$
Bicarbonate-rich group (BG)	W1	Carannante	0.20	-0.03	-0.17	18.1	14.2	12.4
	W2	Castello di Baia	0.26	0.05	-0.31	14.6	11.8	8.3
	W3	Conte	1.08	-0.11	-0.97	11.3	3.4	1.4
	W4	CUS	0.39	-0.42	0.03	4.8	2.2	3.4
	W5	Di Dio	0.91	-0.63	-0.28	11.3	2.4	3.4
	W6	Hotel Neronensis	1.67	0.17	-1.84	7.0	1.6	0.2
	W7	INVITALIA–4	0.74	-0.40	-0.34	7.0	2.2	2.4
	W8	Kyme marmi	0.57	0.14	-0.71	6.2	4.0	1.7
	W9	M. Oltremare	0.47	-0.25	-0.23	5.7	2.8	2.8
	W10	Serapeo (spring)	0.56	0.14	-0.70	17.6	11.5	5.0
	W11	Sud Cantieri	0.52	-0.24	-0.28	6.2	2.9	2.8
	W12	Tortorelli	0.91	-0.93	0.02	29.7	4.7	12.2
Chlorine-rich group (ClG)	W13	Agriturismo fondi di Baia	-0.14	0.26	-0.13	4.0	5.9	4.0
	W14	Centro Ippico	0.19	0.33	-0.51	8.3	9.5	4.1
	W15	De Pisis	0.07	1.11	-1.18	24.7	69.6	7.1
	W16	Esposito	0.14	0.32	-0.46	8.2	9.7	4.5
	W17	Franco Lo Scoglio	-1.50	2.01	-0.51	15.0	498.2	39.9
	W18	Franco Lo Scoglio (spring)	-1.57	2.03	-0.46	13.8	508.6	42.1
	W19	Grotta dell'Acqua (spring)	-0.37	2.04	-1.67	13.1	147.5	3.6
	W20	INVITALIA–15	0.33	0.39	-0.73	7.2	7.7	2.5
	W21	Pozzanghera Terme di Agnano	0.47	0.69	-1.16	14.6	18.1	2.8
	W22	Stufe di Nerone–1	-1.59	2.54	-0.96	5.1	319.2	9.6
	W23	Stufe di Nerone–2	-1.52	2.38	-0.85	6.9	341.4	13.5
	W24	Tempio di Venere	-0.39	1.03	-0.64	5.9	24.4	4.6
	W25	Terme di Baia	-0.80	2.12	-1.32	3.7	68.2	2.2
Sulfate-rich group (SG)	W26	INVITALIA–30	-0.55	-0.10	0.65	2.0	3.1	6.6

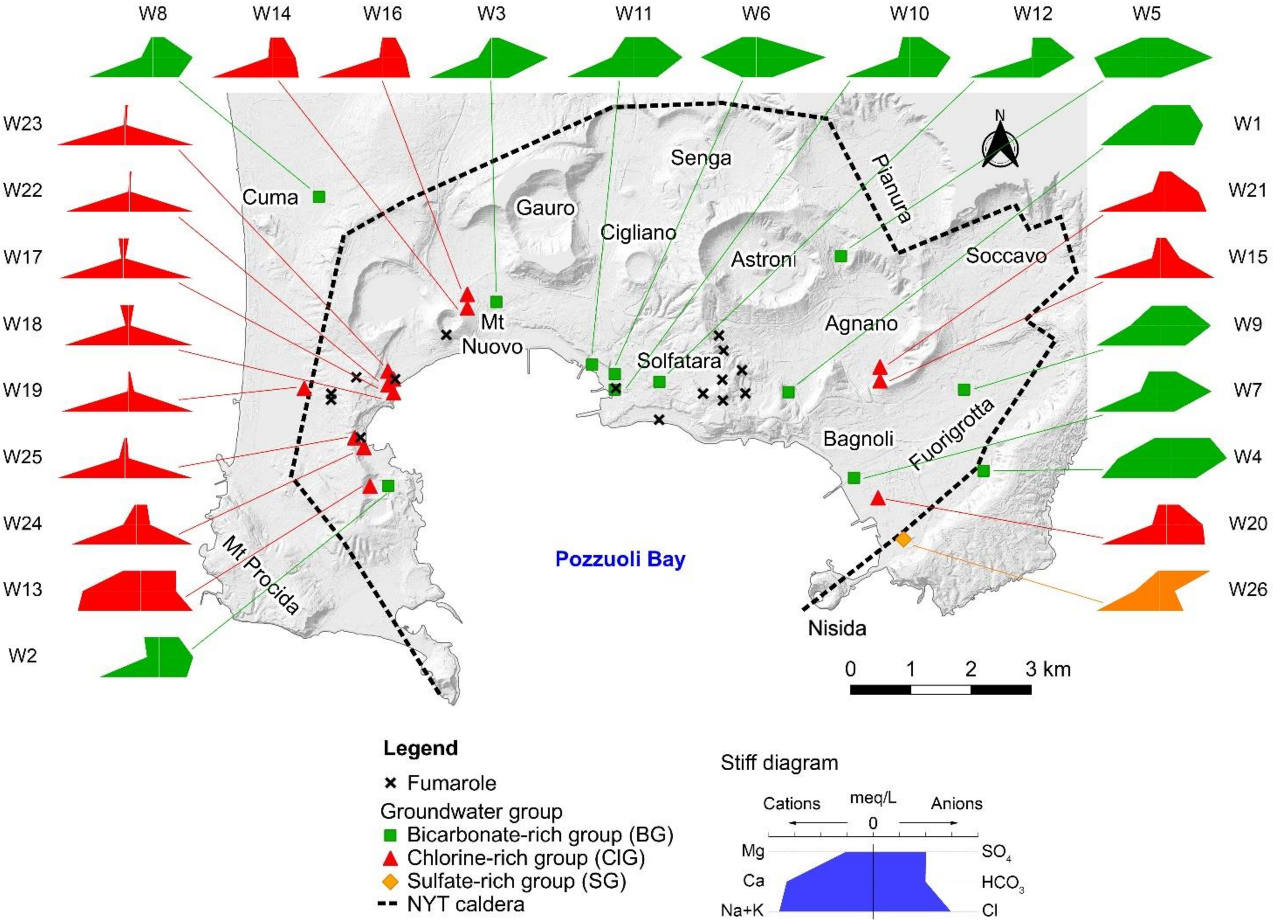


Fig. 2 Visualizing chemical composition of each groundwater sample using the Stiff diagram and spatial distribution of the groundwater groups.

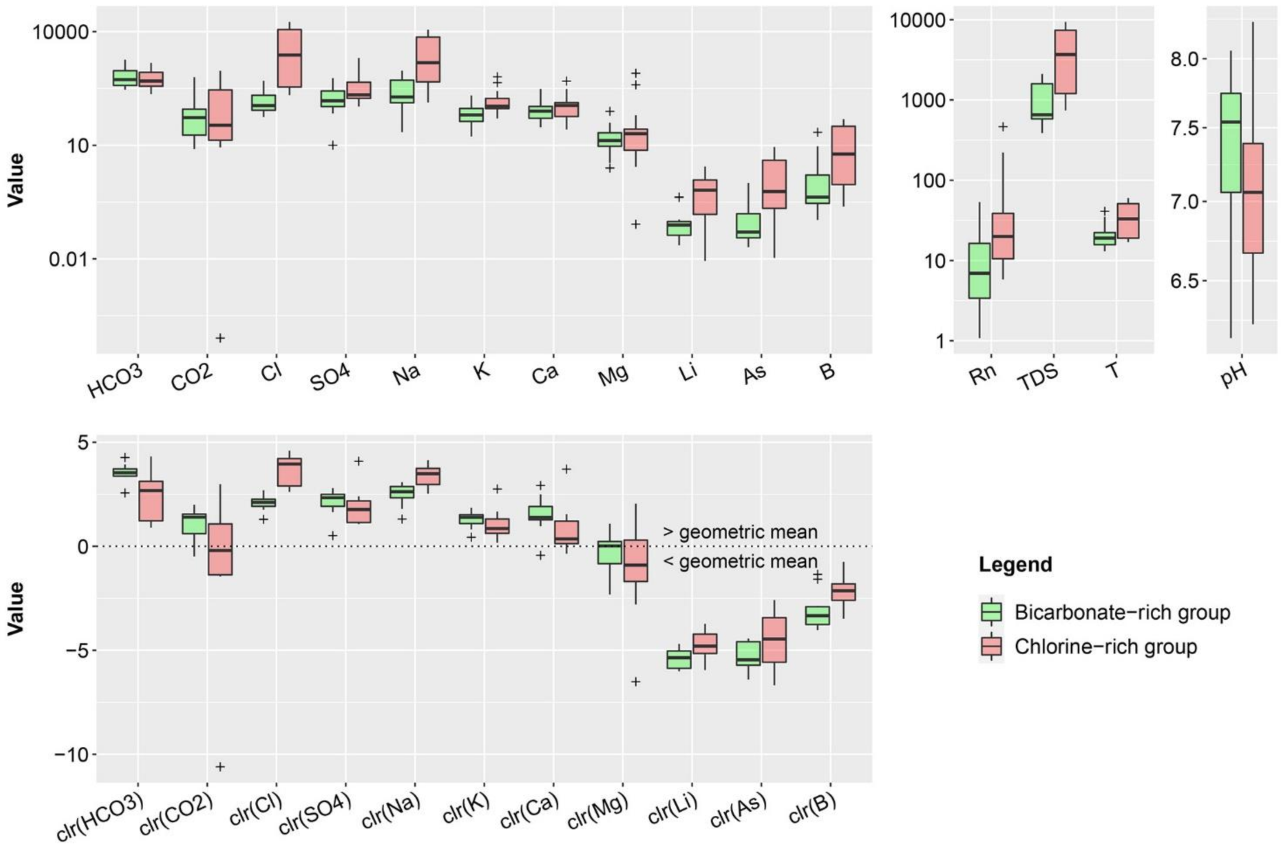


Fig. 3 Chemical composition and physicochemical properties of the bicarbonate- and chlorine-rich groups. The clr-transformation is applied before generating the boxplots in the lower row.

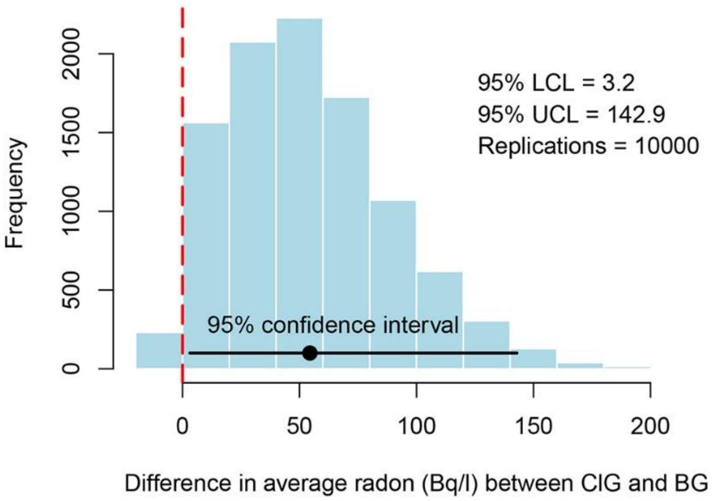


Fig. 4 Distribution of the difference in average radon concentration between the bicarbonate- and chlorine-rich groups. The dot indicates the value calculated from the data in Table 1 and the horizontal line indicates the uncertainty based on the 95% confidence interval of 10000 replications. In each bootstrap replication, average radon value in the bicarbonate-rich group is subtracted from that of the chlorine-rich group. Therefore, the difference is statistically significant if the middle 95% of the distribution is different from zero (i.e. the vertical dashed line).

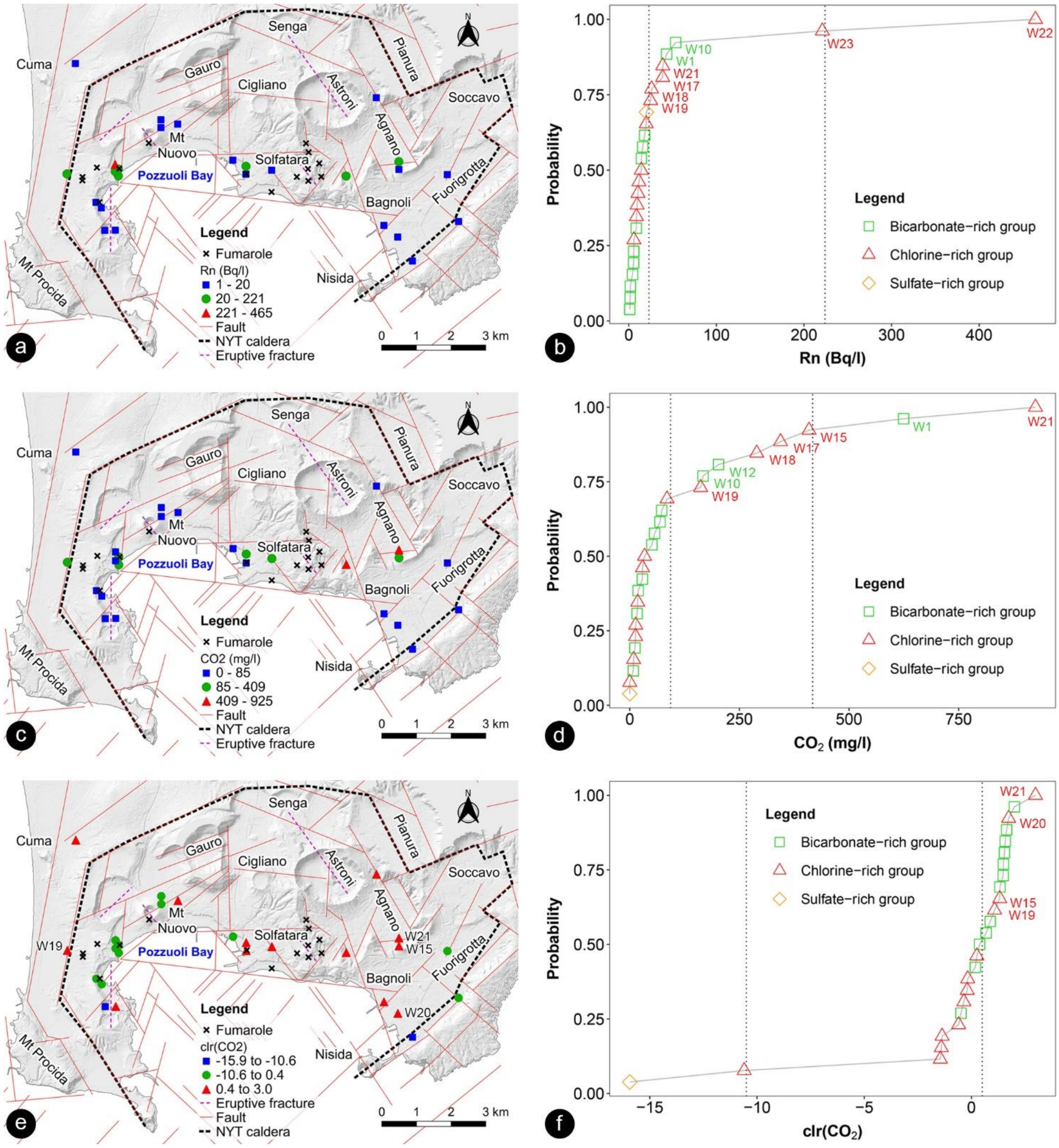


Fig. 5 (a) Spatial distribution of radon in groundwater; (b) The ECDF-plot for radon; (c) Spatial distribution of carbon dioxide in groundwater; (d) The ECDF-plot for carbon dioxide; (e) Spatial distribution of clr-transformed carbon dioxide values in groundwater; and (f) The ECDF-plot for the clr-transformed carbon dioxide. The vertical dotted lines in the ECDF-plots refer to the thresholds (identified by the "natural breaks" method) for generating the dot maps in QGIS.

Radon (Fig. 5a and b) and CO_2 (Fig. 5c and d), whose values are classified via a procedure known as "natural breaks" that identifies the gaps in ordered data, show almost the same spatial distribution in the groundwater aquifer. Although Rn concentration reaches up to 465 Bq/L in the western sector of the study area, CO_2 exceeds 409 mg/L in Agnano plain, where carbon dioxide is one of the main components of the fumaroles. In Fig. 5e and f, considering compositional nature of CO_2 helps to discriminate the bicarbonate-rich group in east Phlegrean Fields from the chlorine-rich group in the west ($clr(CO_2) > 0.4$ versus $clr(CO_2) <$

0.4). The chlorine-rich samples with clr(CO_2) above 0.4 such as De Pisis and Pozzanghera Terme di Agnano (W15 and W21, respectively; Fig. 5e and f) are probably influenced by the carbon dioxide of magmatic origin.

4.2. The main factors regulating groundwater composition

The relationship between temperature and arsenic concentration is revealed in Fig. 6. In the context of compositional data analysis, arsenic balance explains about 50% of variation in temperature ($p < 0.01$; Fig. 6a). However, replacing arsenic balance with natural log of As increases the R-squared to approximately 80% due to spurious correlation ($p < 0.01$; Fig. 6b). Compositional data analysis is therefore coupled with PCA to perform multivariate data analysis and avoid faulty results in this study. Bicarbonate and CO_2 are merged together (amalgamated) prior to conducting PCA because CO_2 is only correlated with bicarbonate probably due to dissolution of the ascending magmatic CO_2 in groundwater and increasing the bicarbonate level. Radon, TDS, pH and temperature were considered as external variables. The results indicate that the first two principal components retain 72.59% of variance (Fig. 7a). There are clearly two associations in the PCA biplot: (1) $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{K^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{HCO_3^{-}}$+$\mathrm{CO_2}$ and pH; and (2) $\mathrm{Na^{+}}$, $\mathrm{Cl^{-}}$, As, B, Li, Rn, TDS and T. The first one is probably related to the groundwater samples originated from meteoric water or contribution of the CO_2-rich magmatic gas in groundwater geochemistry. The most diluted bicarbonate-rich groundwater in north Phlegrean Fields represents a pure meteoric/organic origin of dissolved carbon (Aiuppa et al., 2006). However, the second association expresses the chemical characteristics of hydrothermal fluids and/or seawater. The hot brines (250 to 337 °C) in the deep geothermal reservoir of Phlegrean Fields have considerable concentrations of Li, B and As (13 to 28, 110 to 140 and 9 to 11 mg/L, respectively; Guglielminetti, 1986). Aiuppa et al. (2006) reported a positive correlation between arsenic, chlorine and temperature in the study area which increases from bicarbonate- to chlorine-rich groundwater. They linked the prolonged water–rock interactions (at temperature, $\mathrm{fO_2}$ and $\mathrm{fH_2}S$ of reservoir) under the pH-buffering action of arsenopyrite + pyrite + pyrrhotite mineral assemblage to the elevated As concentrations in NaCl brines. The anomalous Rn quantities could be explained by disintegration of the nonvolatile ^{226}Ra in the deep degassing reservoir (Terray et al., 2018) and in the groundwater aquifer. The measurements in this study are representative of both and correlation of Rn with the variables indicative of hydrothermal fluids and/or seawater may reveal that: (1) the dissolved radon and radium contents are transported by the hydrothermal system of Phlegraean Fields from depth; and (2) the ascending hydrothermal fluids increase groundwater temperature, enhance the interaction between groundwater and aquifer materials, and release ^{226}Ra and consequently Rn in the groundwater body. It is noteworthy that distance of the observations from origin of the biplot demonstrates the Mahalanobis distance of the data point from the mean, being useful for outlier detection (Filzmoser et al., 2018). The samples that lie at the top right corner (i.e. the water from Franco Lo Scoglio and Stufe di Nerone) and the one in the bottom right corner (i.e. the water from Grotta dell'Acqua) of Fig. 7a are likely data outliers. These samples were collected in the Mofete-Fusaro area in west Phlegrean Fields, where geothermal exploration revealed presence of sustained fluxes of deep fluids (Avino et al., 1999). Dilution of the ascending hot NaCl brines by As-depleted meteoric-derived bicarbonate groundwater results in a wide range of As contents in the groundwater body (Aiuppa et al., 2006; Ebrahimi et al., 2022). The groundwater samples with positive PC1 scores (Fig. 7b) are located along the coast or associated with fumaroles, revealing that interaction of the meteoric water with hydrothermal fluids and/or seawater mainly govern hydrogeochemistry of groundwater as flowing towards sea. The results of PCA are in agreement with the background knowledge about the study area, but this analysis is sensitive to outliers and robust estimation of the covariance matrix is impossible due to singularity problem of the clr-transformed values. According to our best knowledge, some research projects will be launched soon to comprehensively characterize radon in the study area and it is therefore recommended to re-evaluate the statistical relationships between variables with a larger dataset and robust methods.

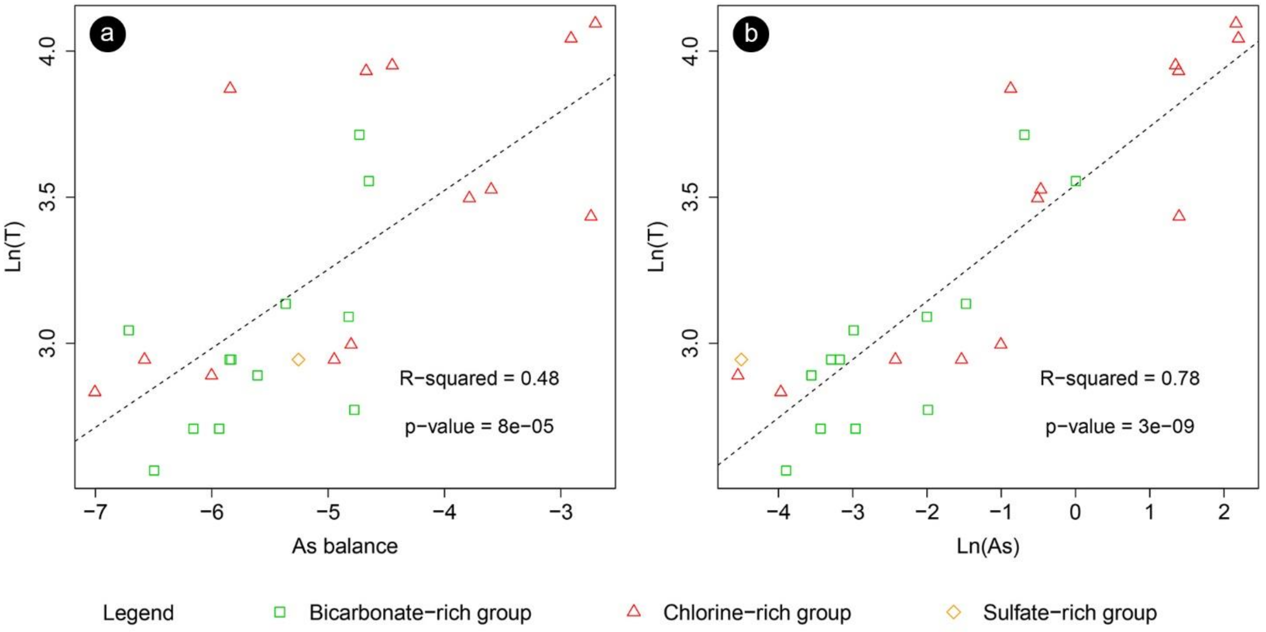


Fig. 6 The relationship between natural log of temperature and: (a) arsenic balance, computed after sequential binary partitioning (Egozcue and Pawlowsky-Glahn, 2005); and (b) natural log of arsenic.

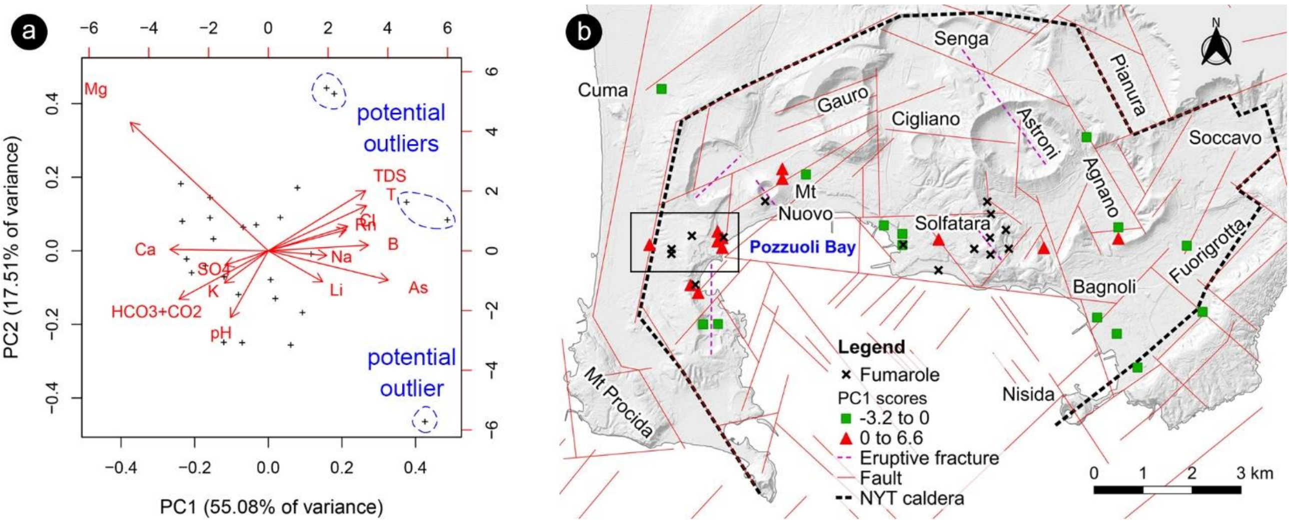


Fig. 7 (a) Principal component analysis biplot in which the first two PCs explain ~73% of variance; and (b) PC1 score map of the groundwater samples. The black box shows the potential outliers in the Mofete-Fusaro area.

5. Conclusions and future work

Monitoring radon and carbon dioxide concentrations is getting more and more popular as a proxy for volcanic activity. Although CO_2 has been studied in Phlegrean Fields in the last decades, Rn measurement has not been a regular practice, especially in underground water. In this study, the dissolved gases (Rn and CO_2), major ions, Li, As, B and other physicochemical parameters are determined in 26 groundwater samples from the central part of Phlegraean Fields caldera to encourage filling the knowledge gap. The samples are classified into bicarbonate-, chlorine- and sulfate-rich groups based on the dominant anion species. Most of the bicarbonate- and chlorine-rich groundwater samples are spatially located in the eastern and western sectors of the study area, respectively. Sulfate is the predominant anion in a sample in the southeast. Spatial distribution of Rn and CO_2 highlights the areas with significant contribution of the hydrothermal system. The highest radon levels are measured in the west, but the elevated carbon dioxide values occur in the east. Dealing with the compositional nature of geochemical data facilitates characterizing the bicarbonate-rich groundwater samples using the clr(CO_2) values. The chlorine-rich group is enriched in $\mathrm{Na^+}$, $\mathrm{Cl^-}$, Li, As and B. The variables are correlated with TDS and temperature in PCA and reveal upflow of hydrothermal fluid and/or seawater through the fault and fracture system, particularly in the Mofete-Fusaro area. On the other hand, $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{K^+}$, $\mathrm{SO_4^{2-}}$, $\mathrm{HCO_3^- + CO_2}$ and pH form an association and highlight the geochemical signature of meteoric water which is influenced by the CO_2-rich magmatic gas in some cases.

In the future investigations, collecting more samples, determining a wider range of chemical variables (including Ra, U and Th), considering the principles of compositional data analysis and performing robust statistical analysis are recommended. Dealing with constant sum constraint might provide new possibilities for monitoring volcanic activity through studying groundwater hydrogeochemistry.

Contribution of the Ph.D. candidate

Collecting samples; Instrumental analysis; Determining research objectives; Data visualization; Data interpretation; Writing the original draft; Revising the draft

CHAPTER 3

Tap Water Chemistry in Naples

This chapter was published in “Environmental Science: Water Research & Technology” as a Research Article:

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. https://doi.org/10.1039/d0ew01123a.

* Corresponding author: stefano.albanese@unina.it

Tap Water Chemistry in Naples

Abstract

Providing safe tap water has been a global concern. Water scarcity, the ever-increasing water demand, temporal variation of water consumption, aging urban water infrastructure and anthropogenic pressure on the water resources are the greatest challenges in effective water supply. In the present chapter, the waters exploited to be introduced in a water distribution system (i.e. input water) and tap waters are collected for determination of metal(loid)s, ions and physicochemical parameters. Seasonal variation is observed in the chemistry of the input waters. Further, the annual total dissolved solids (TDS) of the tap waters range from 200 to 1000 mg/L which stresses the importance of interconnections between urban water reservoirs for mixing different water types and adjusting water quality. It is complicated in populated cities like Naples with an old water distribution network, which also challenges setting up hydraulic and water quality models. The preliminary data visualization indicates the different chemical characteristics of some samples that are supposed to originate from the same source. This might explain the difficulties in understanding the network layout in Naples. Thus, the compositional nature of chemical data was considered in hierarchical cluster analysis (HCA) to seasonally study water transfer between urban water reservoirs and define the source of tap water in each city area. The proposed method can preliminarily divide the pipe network into unique clusters and provide an overview of the relationships between different components when representative models cannot be set up due to limited information about network characteristics. Hence, advanced water distribution simulation and management is encouraged.

1. Introduction

A water distribution system (WDS) is a network of water supply in which consumers, sources and tanks can be the vertices of the network while the connecting pipes, pumps and valves can be the edges (Perelman and Ostfeld, 2011). The size of WDSs has increased over time due to urbanization. It can reach up to hundreds to thousands of nodes and links which substantially enhances the topological (e.g., organization of nodes and links in different combinations) complexities. Further, WDSs are nonlinear dynamic systems and noticeable uncertainties exist in system behaviors, such as the stochasticity of the flow rate, pressure and water demand (Farmani and Butler, 2014).

Nowadays, computer programs (e.g., EPANET and EPANET-PMX; Rossman, 2000; Seyoum and Tanyimboh, 2017) can perform advanced hydraulic and water quality modelling for effective planning and controlling pipe networks along with monitoring water quality within WDSs. Nonetheless, setting up such models is not quite simple because the characteristics of many components with a wide range of properties are needed and many water companies have a vague global vision towards their production and distribution (Izquierdo et al., 2009). Modern guidelines recommend dividing the pipe network into unique clusters called district metered areas (DMAs) before further assessment (Scibetta et al., 2014). Several techniques have been introduced to achieve this aim. Bartolín et al. (2005) constructed a minimum spanning tree to check the unconnected parts of a network, segmentation, broken pipes and conveyance capacity. Likewise, Tzatchkov et al. (2006) used depth-first and breadth-first based graph algorithms for identifying independent networks together with the errors in network data and applying WDS segmentation accordingly. In a generalized graph decomposition model, a network can be characterized by forests (tree structure) and cores (looped structure) as the two main elements which facilitate WDS analysis (Deuerlein, 2008). The network's tree structure was then separated from its looped core and the size of nonlinearity significantly reduced in the upgraded version of the model (Deuerlein et al., 2016). Sempewo et al. (2008) and Di Nardo and Di Natale (2010) presented the potential of spatial analysis for supporting metered area design in WDSs. Similarly, Izquierdo et al. (2009) and Herrera et al. (2010) considered graphical and vector information to design partitioning methods based on machine learning techniques. The clustering coefficient in graph theory can be an indicator of path redundancy in WDSs (Yazdani and Jeffrey, 2011). Perelman and Ostfeld (2011) introduced topological clustering tools to divide a WDS regarding flow direction in pipes. Complex network clustering is another method for the same task (Diao et al., 2013; Scibetta et al., 2013). The modularity-based approach (Diao et al., 2013) was modified by Giustolisi and Ridolfi (2014) through applying a modularity index in multiobjective optimization to obtain various decomposition results for a WDS. Diao et al. (2014) considered the hypothesis that WDSs have been expanded community-by-community during urban development and detected the sub-networks that serve the corresponding communities.

DMAs were introduced in Italy at the end of the 20th century (D.M. 99/97), but they have not widely been used probably due to the complexity of DMA design operations and lack of the required information about the WDS (Giugni et al., 2008). Following many maintenance operations in 2005, Giugni et al. (2008) revealed 67% water loss as a consequence of corroded steel pipes and high-pressure patterns in the eastern sector of Naples (i.e. “Napoli Est”). Accordingly, they considered this part of the city as a DMA and performed hydraulic modelling of the water network using EPANET. The best performance was achieved in a scenario proposing six DMAs by placing 14 intercepting valves together with 9 pressure reducing valves. This solution can save approximately 34% of the water loss. Water consumption patterns were also investigated using different clustering techniques at NW Naples (Padulano et al., 2018).

Controlling the chemical characteristics of tap water is a common safety practice around the world. The companies in charge of managing the water networks annually carry out many chemical analyses with the aim of monitoring water quality and verifying its compatibility with national or local guidelines. The most advanced regulations usually determine the representative analytical parameters and the minimum number of measurements that must be carried out during a year. In Italy, a national legislative decree (Italian Legislative Decree 31/2001) is considered as a reference for monitoring tap water quality. Basically, routine and extended checks need to be performed by the companies distributing water. The former is based on determining a restricted number of parameters in specific checkpoints previously confirmed with the local health authority. The measurements are performed on a seasonal to a daily basis and the parameters include those reported by the European Union (EU Directive 98/83/EC). More stringent criteria are sometimes chosen for the precautionary principle. Furthermore, to ensure a higher level of health protection, the decree recommends monitoring additional parameters to better assess the characteristics of local reservoirs and the status of the hydro-drinking water distribution systems spanning the national territory. For example, vanadium and chlorites are monitored for sanitary reasons while residual disinfectant and hardness are measured as water quality indicators. Although huge analytical databases exist and constantly grow, they have not been used for any other purposes than “to protect human health from the adverse effects of any contamination of water intended for human consumption by ensuring that it is wholesome and clean” (as reported by the European Directive; EU Directive 98/83/EC). These data can be potentially used to study the variations in the amount of water proceeding from different sources. Moreover, the potential influence of construction materials of various infrastructures on the water chemistry can be investigated.

The abovementioned clustering techniques have introduced substantially useful insights towards WDS management. However, they have not been used to improve the efficiency of the old WDS in Naples, one of the most populated cities with a long history in Italy. The studies regarding WDS management are limited in the literature and they do not address the situation in the entire system. It might explain the complex WDS and lack of reliable information about characteristics of the network. Therefore, the optimum operation of the WDS is questionable in the study area. The major aim of this research is to integrate multivariate statistics and compositional data analysis (CoDA) to better understand the WDS of Naples based on the geochemical fingerprints of tap waters. The results of this investigation are useful for partitioning the distribution system into unique clusters and obtaining a general overview about the WDS. It can be a decision support tool for developing hydraulic zones and alleviating the ongoing problems in water distribution operations.

2. Characteristics of the regional and urban water distribution systems

Acqua Bene Comune (ABC) Napoli is a company in charge of providing water services to over two million inhabitants in the city of Naples. The company supplies water to various third-party operators and municipalities; it serves about 280000 delivery points to end users and supplies a daily average water volume of 350000 m^3. The distribution system is based on more than 2000 km of distribution network.

The aqueduct of Naples was inaugurated on 10th May 1885. At the beginning, it was totally fed by the waters of the Serino Aqueduct proceeding from an area, 41 km east of the city, where the Acquaro-Pelosi and Urciuoli springs are located (Fig. 1a). The springs are fed by the carbonate reservoir of the Mt. Terminio massif (Civita, 1969) which belongs to the Apennine chain, a sequence of nappes (alternating siliciclastic and carbonate deposits) overthrusted towards N-NE. Specifically, the water is transported with a pressure line and a free surface covered channel (made of tuff, limestone and brick masonry) for about 60 km from the springs to a receiving tank at the top of the Cancello Hill. From the latter, pressure pipes depart to feed the urban water system and the Lufrano plant (checkpoint).

After the war period (1947-1948), exploitation of the Lufrano aquifer using a complex system of wells began to provide a reserve water supply (AIR) and meet the water demand in case of water scarcity or increasing drinking water consumption (Fig. 1a). The well field has been expanded until the 80s and the water reached the Capodimonte reservoir by a lifting station in Lufrano. The Lufrano aquifer is situated in the southern sector of the Campania Plain, a wide coastal area corresponding to a graben formed in the Upper Pliocene due to the sinking Mesozoic carbonates, and subsequently filled by quaternary alluvial, marine and pyroclastic deposits. This aquifer is fed by the water flowing from the carbonates of Mts. Avella in the east. Thus, Mg^{2+} and Ca^{2+} vs. Na^{+} and K^{+} tends to decrease in groundwater as it flows from the recharge area towards Lufrano (Celico et al., 1994).

The Campano Aqueduct was also joined to the water network in 1958 to meet the increasing water demand in the eastern sector of the city (Fig. 1a). The water supplied by this aqueduct is exploited from the Biferno spring and from the Torano and Maretto springs on the northeastern side and Tyrrhenian side of the Mt. Matese carbonate massif in the Apennines chain, respectively. The well field of Ponte Tavano which is supplied by the carbonate reservoirs of Mts. Durazzano and located a few kilometers north of the Cancello Hill belongs to the latter aqueduct as well.

The Western Campania Aqueduct (ACO) has contributed to the water balance of the city supply system since 1998. At the slopes of a rocky limestone hill in the Lazio sub-Apennines, the water in ACO originates from the springs of the Gari river in Cassino (Latium). The water from Cassino is mixed with the water proceeding from several well fields (i.e. Peccia, Monte Maggiore, S. Sofia, Monte Tifata and S. Prisco) together with the water of the S. Bartolomeo and Sammucro springs and it is transported to the slopes of the S. Prisco Hill, northwest of the city of Caserta (Fig. 1a). The water that reaches the S. Prisco node is divided into four main large pipes that feed the aqueduct of Naples and the Campano Aqueduct storage tank at S. Clemente and Melito.

Between 2007 and 2009, the Cancello well field was also established by ARIN (the former ABC Napoli) in the municipality of S. Felice a Cancello to ensure water balance in the city (Fig. 1a).

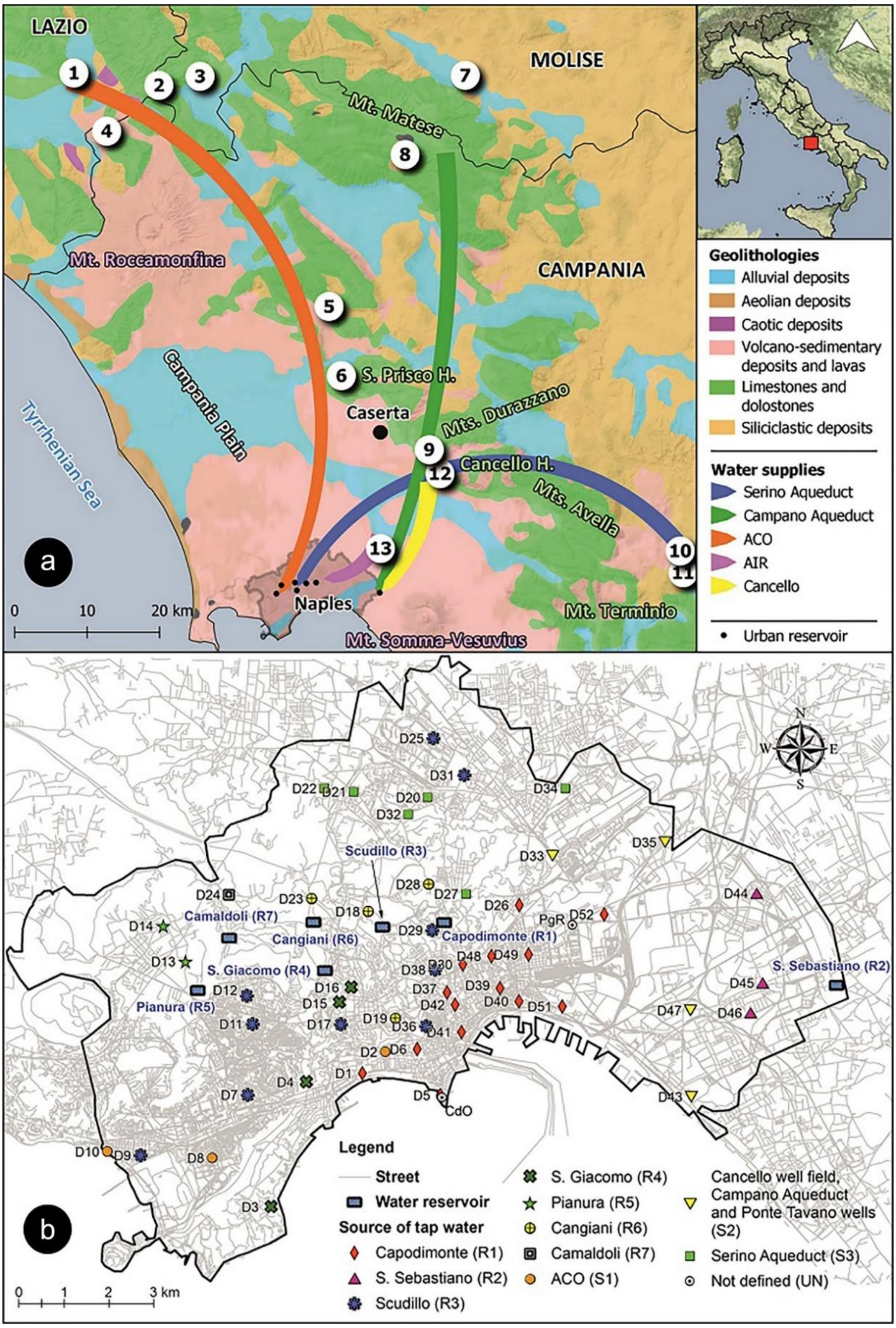


Fig. 1 (a) Lithology of the study area and the regional water supply to the Naples aqueduct. 1: Gari spring; 2: Sammucro spring; 3: S. Bartolomeo spring; 4: Peccia well field; 5: Monte Maggiore well field; 6: S. Prisco, S. Sofia, and Mt. Tifata well fields; 7: Biferno spring; 8: Torano and Maretto springs; 9: Ponte Tavano well field; 10: Urciuoli spring; 11: Acquaro-Pelosi springs; 12: Cancello well field; and 13: Lufrano well field. (b) Location of the collected tap waters (D1 to D52, CdO and PgR) and their source based on Fig. 2.

Before reaching the water network in Naples, the input waters undergo several mixing processes (Fig. 2) which can change the geochemical fingerprint of the natural reservoirs. Moreover, the changes are favored by temporary water storage in a few urban reservoirs constructed by tuff, concrete or a combination of both. Specifically, in the year 2017, the city aqueduct counted on 7 interconnected urban reservoirs (Capodimonte, S. Sebastiano, Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli), with a total storage capacity of approximately 400 million liters, located at different altitudes ranging from sea level to 458 m.a.s.l. at Camaldoli hill.

The Scudillo and S. Giacomo urban reservoirs receive water from the Serino Aqueduct. The S. Giacomo reservoir is also interconnected with the Capodimonte reservoir. The water in the S. Sebastiano reservoir originates from the Cancello well field, Campano Aqueduct and Ponte Tavano wells and, seasonally, from the Serino Aqueduct. The Capodimonte reservoir receives waters from ACO and AIR wells; however, Pianura, Cangiani and Camaldoli reservoirs are connected to the S. Giacomo reservoir (Fig. 2).

The Neapolitan water network has been built in different timespans depending on the expansion of the city and it is consequently possible to find pipelines made of different materials such as gray cast iron, ductile iron, steel, and reinforced concrete. Most of the network developed before the 70s (the historical center and western sector of the city) is characterized by gray cast iron while bituminous steel is present in the recently urbanized areas in east of the city (Poggioreale, Secondigliano, Ponticelli and Barra). Frequent network failures in the latter areas are due to poor conservation.

3. Materials and methods

3.1. Sampling and analysis

The south central region of Italy that overlooks the Tyrrhenian Sea is characterized by dry summers and moderately rainy winters, especially along the coast. Rainfall increases significantly towards the Apennines chain during the cold seasons and the annual average reaches up to 1800 mm (Ducci and Tranfaglia, 2008). Southern Italy, including the study area, experienced an overall reduction of precipitation in the last decades which resulted in decreasing water infiltration and groundwater recharge. In this chapter, the sampling campaigns were conducted in 2017 when winter (including December 2016) and spring were relatively dry. Summer was substantially hot and dry, except for September when the average rainfall was twice the climatological value. Rainfall intensity significantly varied in autumn. October was quite dry, but November was very rainy. Precipitation was in line with the expected climatological values in December.

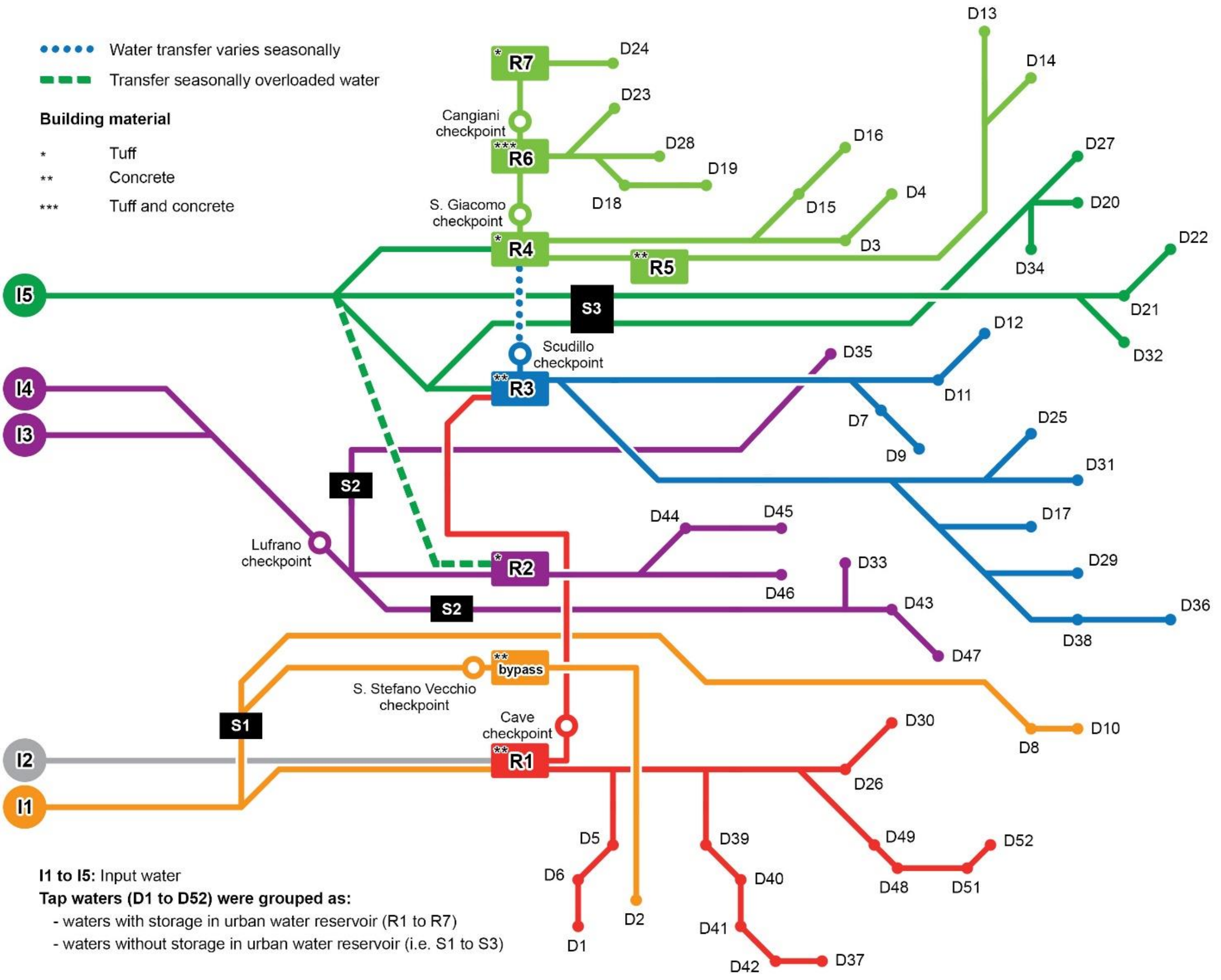


Fig. 2 The tap-water supply scheme of Naples. The input waters are defined as I1: ACO; I2: AIR wells; I3: Cancello well field; I4: Campano Aqueduct and Ponte Tavano wells; and I5: Serino Aqueduct. D1 to D52 represent the tap water samples. Every month, the local water company collects tap water at these stations to check quality and the results are accessible to the public. The coordinates and address of these stations are provided in Table 1. In this study, the tap waters stored in urban water reservoirs are grouped as R1 to R7 and the rest are categorized as S1 to S3. The urban water reservoirs are shown as R1: Capodimonte (92 m.a.s.l.); R2: S. Sebastiano (112 m.a.s.l.); R3: Scudillo (183 m. a.s.l.); R4: S. Giacomo (230 m.a.s.l.); R5: Pianura (220 m.a.s.l.); R6: Cangiani (303 m.a.s.l.); and R7: Camaldoli (458 m.a.s.l.). However, S1, S2 and S3 are defined as follows: S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; and S3: Serino Aqueduct.

To check water suitability for human consumption, NO_3^-, EC and pH were regularly measured (661 determinations for nitrate, but 664 measurements for EC and pH) in the input waters (I1 to I5 in Fig. 2) by qualified personnel of ABC Napoli. Moreover, concentrations of Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-}, Cl^- and F^- were determined monthly in AIR wells (I2) and seasonally in the other input waters. Further information about the number of samples is available in Table 2 and Fig. 3. Tap water samples (Fig. 1b) were monthly collected from 53 sampling points (a total of 626 samples) across Naples after cleaning the bottles. Briefly, the bottles were rinsed several times with ultrapure water. The absence of both "memory" effect and release of chemicals from the container was then verified by filling randomly selected bottles with ultrapure water and checking composition of the content analytically. The procedure has been approved by the Italian sole accreditation body (Accredia). In the sampling campaigns, the tap was turned on for at least 15 minutes (10 minutes more in the case of turbid water) and constant water temperature was attained prior to sampling. The bottles were first rinsed with tap water several times and then filled with water completely. The samples were stored at 5 ± 3 °C and transported to the Laboratorio Controllo Acque at ABC Napoli for chemical analyses. $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{Na^{+}}$, $\mathrm{K^{+}}$, $\mathrm{NH_4^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{Cl^{-}}$, $\mathrm{NO_2^{-}}$, $\mathrm{NO_3^{-}}$, $\mathrm{F^{-}}$ and $\mathrm{ClO_2^{-}}$ were detected using ion chromatography (IC); however, $\mathrm{HCO_3^{-}}$, $\mathrm{Cl_2}$, TDS and turbidity were determined via titration, colorimetry, gravimetric method and nephelometry, respectively. Inductively coupled plasma-mass spectrometry (ICP-MS) was also used for measurement of As, Cu, Pb, Cd, Ni, Cr, V, Hg, Fe, Mn and Al. Determination of electric conductivity (EC) and pH with a Mettler Toledo T70, and temperature with a digital thermometer was done in the field. The following formula which accounts for $\mathrm{Ca^{2+}}$ and $\mathrm{Mg^{2+}}$ as the main factors for total hardness (TH) was also applied for TH computation (Todd, 1980):

$$\mathrm{TH} = 2.5\mathrm{Ca^{2+}} + 4.1\mathrm{Mg^{2+}} \tag{1}$$

where all values were in mg/L.

The detection limits were calculated as three times the standard deviation of the analyte concentration in the samples with very low content of the analyte: 0.01 μg/L for Hg, 0.1 μg/L for As, Pb, Cd, Ni, Cr, V, Mn and Al, 1 μg/L for Cu, Fe and $\mathrm{ClO_2^{-}}$, 0.01 mg/L for $\mathrm{NO_2^{-}}$, 0.02 mg/L for $\mathrm{NH_4^{+}}$ and $\mathrm{Cl_2}$, 0.03 mg/L for $\mathrm{F^{-}}$, 0.07 mg/L for $\mathrm{NO_3^{-}}$, 0.1 mg/L for $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$ and $\mathrm{SO_4^{2-}}$, 0.2 mg/L for $\mathrm{Na^{+}}$, 0.25 mg/L for $\mathrm{Cl^{-}}$, 0.4 mg/L for $\mathrm{K^{+}}$, 2 mg/L for $\mathrm{HCO_3^{-}}$, 10 mg/L for TDS, 1 μS/cm for EC and 0.05 NTU for turbidity. Blanks, trace element matrix reference materials (RMs) and NIST standard reference materials (SRMs) were employed to verify measurement precision and accuracy. The maximum calculated error is 11% and relative standard deviations (% RSDs) of the measurements range between 0 and 7%.

3.2. Data preparation and statistical analysis

Different steps from data treatment to data visualization and statistical analysis are summarized in Fig. 4 and further discussed here. Once the experimental results were available, the samples with unacceptable cation–anion balance (CAB) error (i.e. not within ±10%) were omitted and 577 tap water samples remained. Cu, Pb, Hg, Cd, Cr, $\mathrm{NO_2^{-}}$ and $\mathrm{NH_4^{+}}$ were ignored because most of the values were below the corresponding detection limits (Fig. 4).

Combined with conventional statistical analyses, compositional data analysis is a powerful tool to get better knowledge of the relationships and dependencies between elements, various geochemical processes and different sources of chemicals (Somma et al., 2021; Thiombane et al., 2018a). Thus, the hierarchical clustering algorithm was coupled with CoDA in this chapter. To apply compositional data analysis and statistical analysis, a dataset with no value below the detection limits is necessary. Because the proportion of rounded zeros (the values below the detection limit) for Ni, Fe, Mn, Al and $\mathrm{ClO_2^{-}}$ is higher than 10%, a complete data matrix with nonzero entries was obtained through running a multivariate modified EM algorithm after additive log-ratio transformation of the compositions (Fig. 4) (Palarea-Albaladejo and Martín-Fernández, 2008). This parametric approach consists of modifying the EM-algorithm in order to include the (log-ratio) imputed values with smaller values than the (log-ratio) detection limit. The “zCompositions” package of R software was utilized for zero replacement.

Table 1 Coordinates and address of the stations for regular tap water sampling in Naples.

Station	Latitude	Longitude	Address
D1	40.83293	14.22750	Riviera di Chiaia
D2	40.83725	14.23346	Piazza Amedeo
D3	40.80584	14.20356	Via Posillipo
D4	40.83105	14.21260	Via Manzoni
D5	40.82864	14.24821	Piazzetta Marinari
D6	40.83786	14.24191	Piazza Mondragone
D7	40.82834	14.19706	Via Benedetto Cariteo
D8	40.81558	14.18782	Via Cavalleggeri d'Aosta
D9	40.81598	14.16891	Via S. Ferrara
D10	40.81671	14.16007	Via di Pozzuoli
D11	40.84262	14.19819	Via Adriano
D12	40.84828	14.19665	Via Scherillo
D13	40.85508	14.18035	Via Padula
D14	40.86217	14.17434	Corso Duca d'Aosta
D15	40.84716	14.22119	Via Caldieri
D16	40.85015	14.22429	Via Simone Martini
D17	40.84272	14.22162	Via Santa Maria Della Libera
D18	40.86547	14.22863	Via del Serbatoio
D19	40.84402	14.23613	Via Ligorio Pirro
D20	40.88852	14.24409	Via V. Veneto
D21	40.88952	14.22445	Via E. Scaglione
D22	40.89018	14.21661	Corso Chiaiano
D23	40.86793	14.21358	Via L. Bianchi
D24	40.86862	14.19166	Via Guantai ad Orsolona
D25	40.90044	14.24549	Via A. Ghisleri
D26	40.86703	14.26864	Via Arenaccia
D27	40.86916	14.25448	Via Ponti Rossi
D28	40.87105	14.24460	Viale Colli Aminei
D29	40.86182	14.24565	Vicoletto S. Vincenzo alla Sanità
D30	40.85501	14.25383	Piazza Cavour
D31	40.89304	14.25370	Via Lombardia
D32	40.88509	14.23899	Via G. Manfredi ang. Via Lieti a Capodimonte
D33	40.87699	14.27734	Varco ingresso dogana Aeroporto - Viale U. Maddalena
D34	40.89060	14.28059	Via G. Pascale
D35	40.87987	14.30715	Via Icaro
D36	40.84251	14.24419	Corso V. Emanuele
D37	40.84932	14.24976	Piazza Dante
D38	40.85388	14.24652	Via S. Rosa
D39	40.85024	14.26379	Piazza Calenda
D40	40.84766	14.26880	Corso Garibaldi
D41	40.84137	14.25368	Via Flavio Gioia
D42	40.84685	14.25189	Piazza del Gesu
D43	40.82874	14.31460	Via Bernardino Martirano
D44	40.86975	14.33157	Via Madonnelle
D45	40.85177	14.33326	C.so Ponticelli
D46	40.84571	14.33017	Via Mastellone
D47	40.84610	14.31420	Via Volpicella
D48	40.85672	14.26142	Via C. Rosaroll
D49	40.85710	14.27124	C.ta Ponte di Casanova
D51	40.84670	14.28033	Via Brin
D52	40.86529	14.29121	Largo Tarantini

Table 2 Composition of the input water in the water distribution system of Naples in 2017.

Input	Seasonal/ Annual	Ca^{2+}	Mg^{2+}	Na^{+}	SO_4^{2-}	Cl^{-}	F^{-}	NO_3^{-}	EC	pH
		mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L	µS/cm	-
I1 [a]	Winter	121.6	28.6	8.0	15.0	11.0	0.2	3.9	724.3	7.1
	Spring	124.3	31.0	10.0	18.0	16.0	0.2	4.2	722.9	7.2
	Summer	116.8	28.5	8.0	18.0	12.0	0.2	4.1	736.2	7.2
	Fall	123.8	30.2	8.0	22.0	13.0	0.2	3.8	726.6	7.1
	Annual	121.6	29.6	8.5	18.3	13.0	0.2	4.0	727.5	7.2
I2 [b]	Winter	131.8	27.0	-	-	86.9	1.8	106.3	1172.0	7.2
	Spring	136.8	28.1	80.0	168.0	89.2	1.7	111.1	1165.8	7.3
	Summer	159.8	32.0	-	-	124.9	1.5	84.6	1265.1	7.1
	Fall	153.8	33.5	84.0	152.0	105.1	2.0	104.8	1170.2	7.2
	Annual	145.3	30.2	82.0	160.0	102.0	1.8	101.8	1193.2	7.2
I3 [c]	Winter	165.2	40.4	87.0	46.0	172.0	0.6	9.5	1410.7	6.9
	Spring	185.2	49.4	97.0	52.0	197.0	0.6	9.1	1419.9	7.0
	Summer	190.9	48.2	102.0	51.0	201.0	0.4	9.1	1489.5	6.9
	Fall	172.9	43.1	88.0	43.0	164.0	0.6	9.3	1343.0	6.9
	Annual	178.6	45.3	93.5	48.0	183.5	0.5	9.2	1415.7	6.9
I4 [d]	Winter	188.0	51.8	102.0	59.0	200.0	0.8	13.0	1596.2	7.0
	Spring	-	-	-	-	-	-	13.0	1607.7	7.0
	Summer	-	-	-	-	-	-	12.7	1560.3	7.1
	Fall	202.5	55.6	100.0	58.0	217.0	0.7	13.2	1615.7	6.9
	Annual	195.3	53.7	101.0	58.5	208.5	0.7	13.0	1596.6	7.0
I5 [e]	Winter	54.1	9.6	7.0	9.0	9.0	0.2	4.7	351.2	7.7
	Spring	-	-	-	-	-	-	4.2	347.4	7.8
	Summer	55.0	8.9	7.0	5.0	10.0	0.2	4.0	335.7	7.9
	Fall	-	-	-	-	-	-	4.2	339.8	7.6
	Annual	54.6	9.3	7.0	7.0	9.5	0.2	4.3	345.2	7.8

[a] ACO; [b] AIR wells; [c] Cancello well field; [d] Campano Aqueduct and Ponte Tavano wells; [e] Serino Aqueduct. EC, NO_3^{-} and pH were regularly measured (see Fig. 3 for further details). For the rest of chemicals, results of six measurements in AIR wells and one measurement in the others are available in each season.

The chemicals that feasibly induce adverse effects on the statistical analysis were then culled using univariate statistics (Table 3). To be very concerned with the deleterious effects of measurement errors, the problematic group of chemicals were detected using the proportion of concentrations less than three times the first instrumental calibration point. It was assumed that such low concentrations have noticeable errors. After culling the analytes exceeding the maximum acceptable limit of 10%, Ca^{2+}, SO_4^{2-}, HCO_3^{-}, Cl^{-}, NO_3^{-}, EC and TDS were retained. In addition, pH and temperature that have no detection limits were included. TH was also ignored because of using Mg^{2+} with approximately 19% of values below three times the first instrumental calibration point to calculate the parameter (Table 3 and Fig. 4).

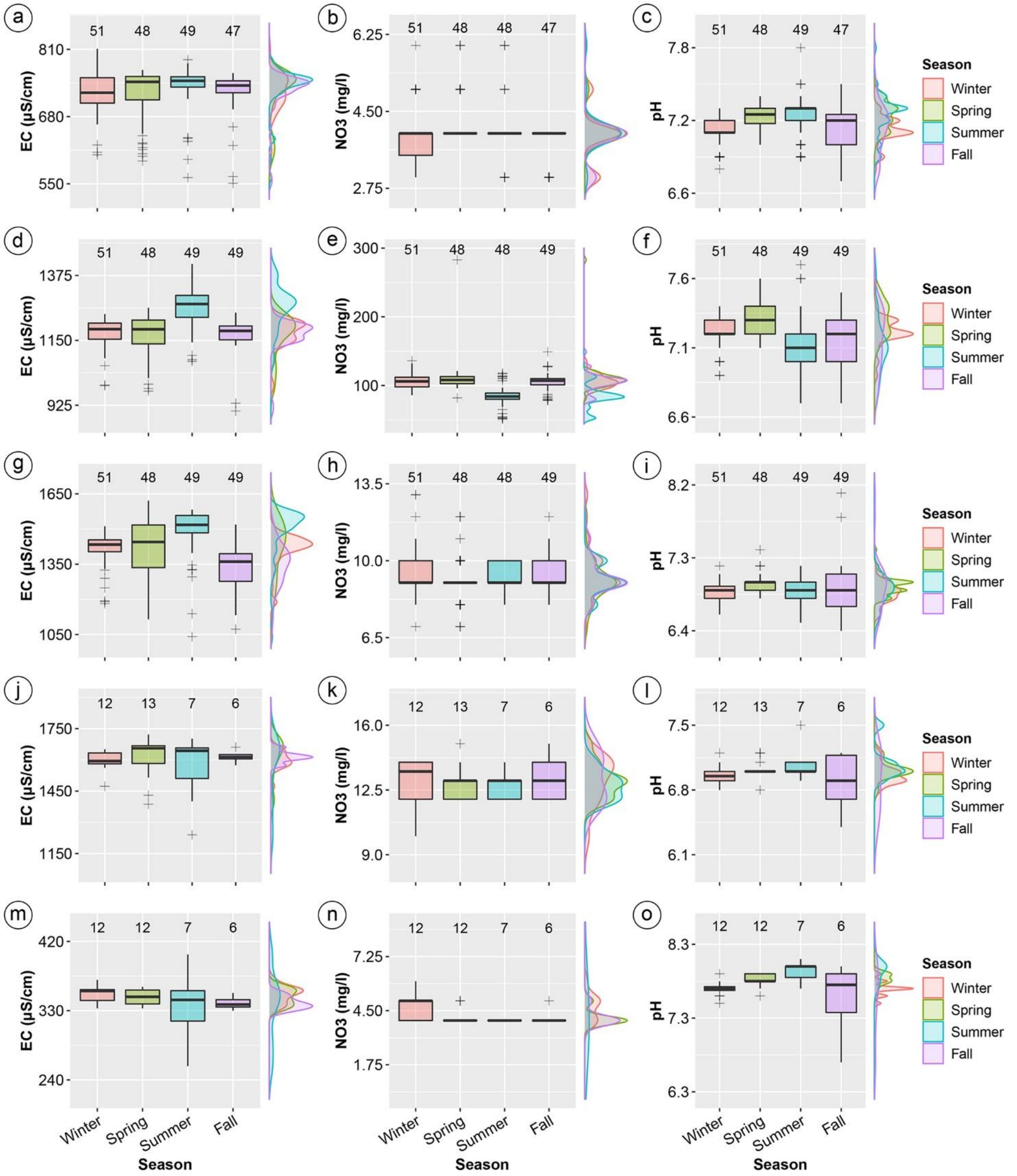


Fig. 3 Distribution of the measured EC, NO_3^- and pH values in the input waters for each season of 2017. The numbers above boxplots show the number of measurements in the corresponding seasons. (a-c) I1: ACO; (d-f) I2: AIR wells; (g-i) I3: Cancello well field; (j-l) I4: Campano Aqueduct and Ponte Tavano wells; (m-o) I5: Serino Aqueduct.

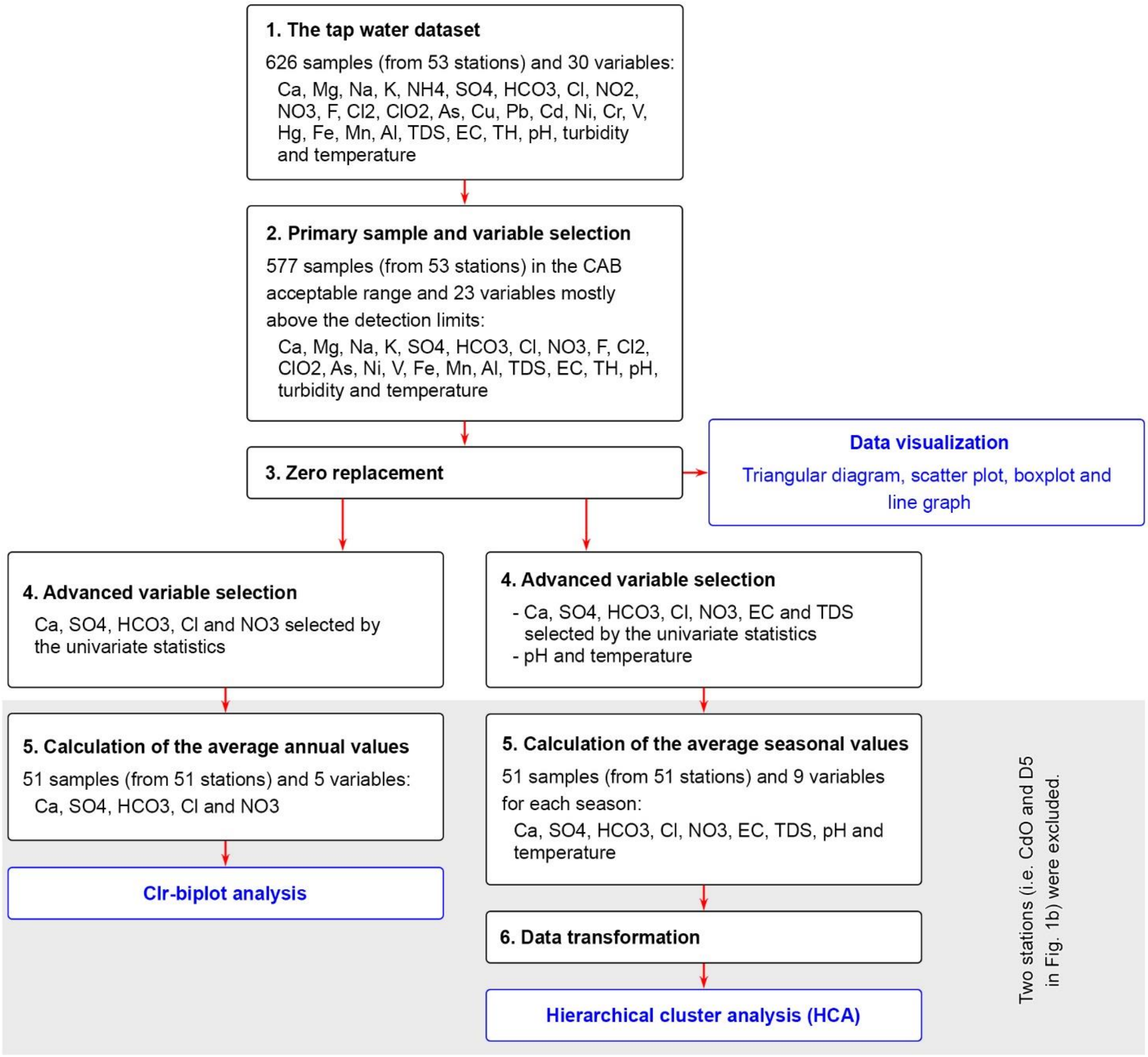


Fig. 4 The flowchart depicting different steps of data treatment before data visualization and statistical analysis.

Because it is expected that precipitation and evaporation influence chemistry of the input water, the seasonal geometric means of the monthly determined variables were subsequently calculated for each station (Fig. 4) to develop this study based on each season. Two sampling stations (i.e. CdO and D5) without analytical results in some seasons were also ignored to keep the same number of samples between seasons. Due to the compositional nature of chemical data, the ratio between chemicals provides more information than the absolute measured values (Pawlowsky-Glahn and Buccianti, 2011). Therefore, the dataset was transformed by means of the centered log-ratio (clr) transformation proposed by Aitchison (1982). The advantage of clr- over alr- (additive log-ratio; Aitchison, 1982) transformation is that it represents each variable as a ratio to a central value (the geometric mean of a subcomposition) instead of favoring one variable (Dth part) in all denominators (Aitchison, 1986). However, this approach can lead to some difficulties in interpretating clr values as the concentration of the original elements because any single component can be influenced by enrichment or depletion in the set of components considered. Nevertheless, clr-transformation was chosen because it is effective for understanding the underlying geochemical processes (McKinley et al., 2016). Specifically, the clr-transformation for a composition $x = [x_1, x_2, \ldots, x_D] \in S^D$ is the transformed data $y \in R^D$ with (Lloyd et al., 2012):

$$y = \mathrm{clr}(x) = \left(\ln \frac{x_1}{g(x)}, \ln \frac{x_2}{g(x)}, \ldots, \ln \frac{x_D}{g(x)}\right) = \ln \frac{x}{g(x)} \tag{2}$$

where g(x) is the geometric mean across the composition x which can be computed as:

$$g(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} = \sqrt[n]{x_1 . x_2 \ldots x_n} = \exp\left[\frac{1}{n}\sum_{i=1}^{n} \ln x_i\right] \tag{3}$$

EC, TDS and temperature that cannot be treated like compositional data were log-transformed and no transformation was applied to pH due to its logarithmic scale (Fig. 4). Finally, for hierarchical cluster analysis (HCA), the Euclidean distance (for computation of the distances between the samples) and Ward's method (for agglomerative clustering) were used as the most widely applied methods that form easily interpretable and possibly significant clusters in the hydrogeochemical, hydrologic and geologic contexts (Fig. 4) (Gong and Richman, 1995; Güler et al., 2002; Keshavarzi et al., 2015). The CoDaPack v2.02.21 and “cluster” R package were used for the data transformation and HCA, respectively.

In order to better understand the compositional variability, the associations between the parts of a (sub)composition and the likely natural sources of tap water, a clr-biplot was also generated using the annual geometric mean of Ca^{2+}, HCO_3^-, Cl^-, SO_4^{2-} and NO_3^- in each station via the CoDaPack v2.02.21. Biplots graphically represent the clr-transformed compositional data based on the principal component analysis (Aitchison, 1983). Construction of a clr-biplot starts with clr transformation of the data matrix via eqn (2). The clr-components are then centered in order to put the origin of axes in the center of the data. In the next step, a singular value decomposition (SVD) of the centered clr-data is conducted to generate orthogonal axes showing maximum variability of the dataset. The observations and the variables are finally demonstrated as dots and rays (vectors) on a plot, respectively (Martín-Fernández and Thió-Henestrosa, 2016). The data points with similar values will be clustered on biplots. The squared length of a clr-ray is comparable with the clr-variance of the corresponding variable. Nevertheless, the squared length of the link between two vertices is almost equal to the variance of the log-ratio of those two parts. Hence, the correlation of two variables increases with decreasing distance between the corresponding vertices (Aitchison and Greenacre, 2002; Zuzolo et al., 2020). Uncorrelated components are at right angles to each other. Further, the cosine of the angle between two intersecting links explains the correlation coefficient between the log-ratios. Consequently, if two links are at right angles, zero correlation of the log-ratios is expected (Pawlowsky-Glahn and Buccianti, 2011).

4. Results and discussion

4.1. Chemistry of the input water

According to Fig. 2, the water from ACO (I1) and AIR wells (I2) is delivered to the Capodimonte reservoir (R1). Ca^{2+}, Na^+, Cl^-, SO_4^{2-}, NO_3^-, F^- and EC values are low in the former and high in the latter which result in their different chemical patterns (Table 2 and Figs. 3, 5 and 6). The highest concentrations of SO_4^{2-}, NO_3^- and F^- are determined in I2; however, I3 (Cancello well field) and I4 (Campano Aqueduct and

Ponte Tavano wells) account for the highest values of the other parameters (Table 2 and Fig. 5). The analytical results of I3 and I4 are almost the same leading to their similar chemical signature in Fig. 5. The lowest chemical contents are reported for the slightly basic waters originating from the Serino Aqueduct (I5). The pH values indicate the neutral characteristics of the other input waters. Statistical distributions of nitrate, EC and pH which were measured regularly over the timespan of this study indicate that composition of the input water probably reflects precipitation and evaporation in the study area (Fig. 3). Mixing different proportions of the waters with high (i.e. I2, I3 and I4) and low (i.e. I1 and I5) dissolved contents plays an important role in meeting the quality standards (Italian Legislative Decree 31/2001). However, changes in volume and composition of the input water during the year can be a great challenge in water transfer between the urban water reservoirs to satisfy the increasing water demand of the inhabitants (Table 2 and Figs. 3 and 6).

4.2. Chemistry of tap water

The chemical characteristics of tap water are statistically summarized in Table 3. Even though the average concentrations of cations are in the decreasing order of $\mathrm{Ca^{2+}}$ (116.5 mg/L) > $\mathrm{Na^{+}}$ (29.3 mg/L) > $\mathrm{Mg^{2+}}$ (27.8 mg/L) > $\mathrm{K^{+}}$ (10.9 mg/L), the mean contents of anions follow the order of $\mathrm{HCO_3^-}$ (472.6 mg/L) > $\mathrm{Cl^-}$ (47.4 mg/L) > $\mathrm{SO_4^{2-}}$ (33.1 mg/L) > $\mathrm{NO_3^-}$ (16.5 mg/L) > $\mathrm{F^-}$ (0.4 mg/L). On average, the selected metal(loid)s (As, Ni, V, Fe, Mn and Al) are below 6 μg/L. Chlorine dioxide ($\mathrm{ClO_2^-}$) and chlorine ($\mathrm{Cl_2}$), which are utilized for chemical disinfection of water, do not exceed 153 μg/L and 0.3 mg/L, respectively, in 95% of the samples. The average and median of $\mathrm{Cl_2}$, $\mathrm{Ca^{2+}}$, $\mathrm{Mg^{2+}}$, $\mathrm{K^{+}}$, $\mathrm{SO_4^{2-}}$, $\mathrm{HCO_3^-}$, $\mathrm{NO_3^-}$, EC, TDS, TH, pH and temperature are approximately equal. Tap waters are colder than 23.5 °C and indicate slightly acidic to basic nature (pH = 6.6 to 8.4). The interquartile range (IQR) explains the variance of the measured parameters. In this chapter, metal(loid)s, $\mathrm{Cl_2}$, $\mathrm{F^-}$, turbidity, pH and temperature have the lowest variation (IQR < 5). Whilst most major ions indicate medium variance (IQR = 5-100), the IQR values of EC, TDS, TH and bicarbonate are the greatest (>100) (Table 3).

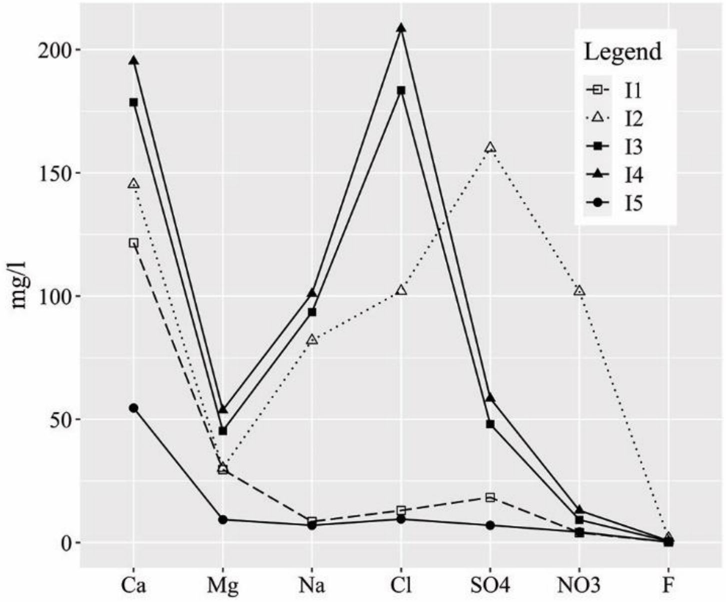


Fig. 5 Chemical compositions of the input waters in the water distribution system of Naples in 2017. I1: ACO; I2: AIR wells; I3: Cancello well field; I4: Campano Aqueduct and Ponte Tavano wells; and I5: Serino Aqueduct.

Table 3 Statistical summary of the measured parameters in the selected 577 tap water samples

Parameter	**FCP** [a]	**%<3FCP**	**Min**	**P5**	**P25**	**P50** [b]	**P75**	**P95**	**Max**	**IQR**	**Average**
As (µg/L)	1.0	42.1	<D.L. [c]	1.0	2.0	3.0	3.0	3.0	4.0	1.0	2.4
Ni (µg/L)	1.0	86.8	<D.L.	0.4	0.6	1.0	1.0	5.0	13.0	0.4	1.4
V (µg/L)	1.0	23.2	1.0	2.0	3.0	3.0	4.0	6.0	8.0	1.0	3.5
Fe (µg/L)	1.0	37.6	<D.L.	0.6	2.0	4.0	6.0	18.0	122.0	4.0	5.9
Mn (µg/L)	1.0	58.1	<D.L.	0.5	1.0	2.0	4.0	9.2	29.0	3.0	3.1
Al (µg/L)	1.0	85.1	<D.L.	0.3	0.4	1.0	2.0	5.0	140.0	1.6	2.0
$\mathrm{ClO_2^-}$ (µg/L)	50.0	94.3	<D.L.	1.6	4.1	6.6	71.0	152.6	312.0	66.9	39.3
$\mathrm{Cl_2}$ (mg/L)	0.08	84.2	<D.L.	0.11	0.16	0.2	0.22	0.26	0.44	0.06	0.2
$\mathrm{Ca^{2+}}$ (mg/L)	12.0	0.0	46.4	56.4	100.9	121.3	132.9	180.3	204.1	32.0	116.5
$\mathrm{Mg^{2+}}$ (mg/L)	6.0	18.5	6.9	9.6	24.1	30.1	32.1	45.0	49.9	8.0	27.8
$\mathrm{Na^+}$ (mg/L)	6.0	36.7	4.0	7.0	13.0	23.0	36.0	93.0	108.0	23.0	29.3
$\mathrm{K^+}$ (mg/L)	3.0	36.2	1.0	3.0	6.3	11.6	14.4	18.9	23.9	8.1	10.9
$\mathrm{SO_4^{2-}}$ (mg/L)	1.0	0.7	1.0	9.0	19.0	36.0	46.0	55.2	64.0	27.0	33.1
$\mathrm{HCO_3^-}$ (mg/L)	2.0	0.0	183.0	221.6	411.0	505.0	534.0	698.0	754.0	123.0	472.6
$\mathrm{Cl^-}$ (mg/L)	0.5	0.0	8.0	9.0	16.0	31.0	53.0	184.0	219.0	37.0	47.4
$\mathrm{F^-}$ (mg/L)	0.09	27.9	0.1	0.2	0.2	0.5	0.6	0.7	0.9	0.3	0.4
$\mathrm{NO_3^-}$ (mg/L)	1.0	0.3	<D.L.	4.0	8.0	17.0	22.0	32.0	38.0	14.0	16.5
EC (µS/cm)	1.0	0.0	259.0	345.0	651.0	798.0	916.0	1405.6	1610.0	265.0	789.5
TDS (mg/L)	10.0	0.0	181.0	242.0	456.0	559.0	641.0	984.4	1127.0	185.0	552.7
TH (mg/L)	-	-	149.8	180.5	361.6	425.0	462.9	632.0	713.6	101.3	405.1
Turbidity (NTU)	0.05	24.6	<D.L.	0.1	0.2	0.2	0.3	0.4	1.1	0.1	0.2
pH	-	-	6.6	6.8	7.1	7.3	7.5	7.9	8.4	0.4	7.3
T (°C)	-	-	11.1	11.9	13.7	14.4	15.0	17.0	23.5	1.3	14.4

[a] The first instrumental calibration point. [b] Median. [c] < Detection limit.

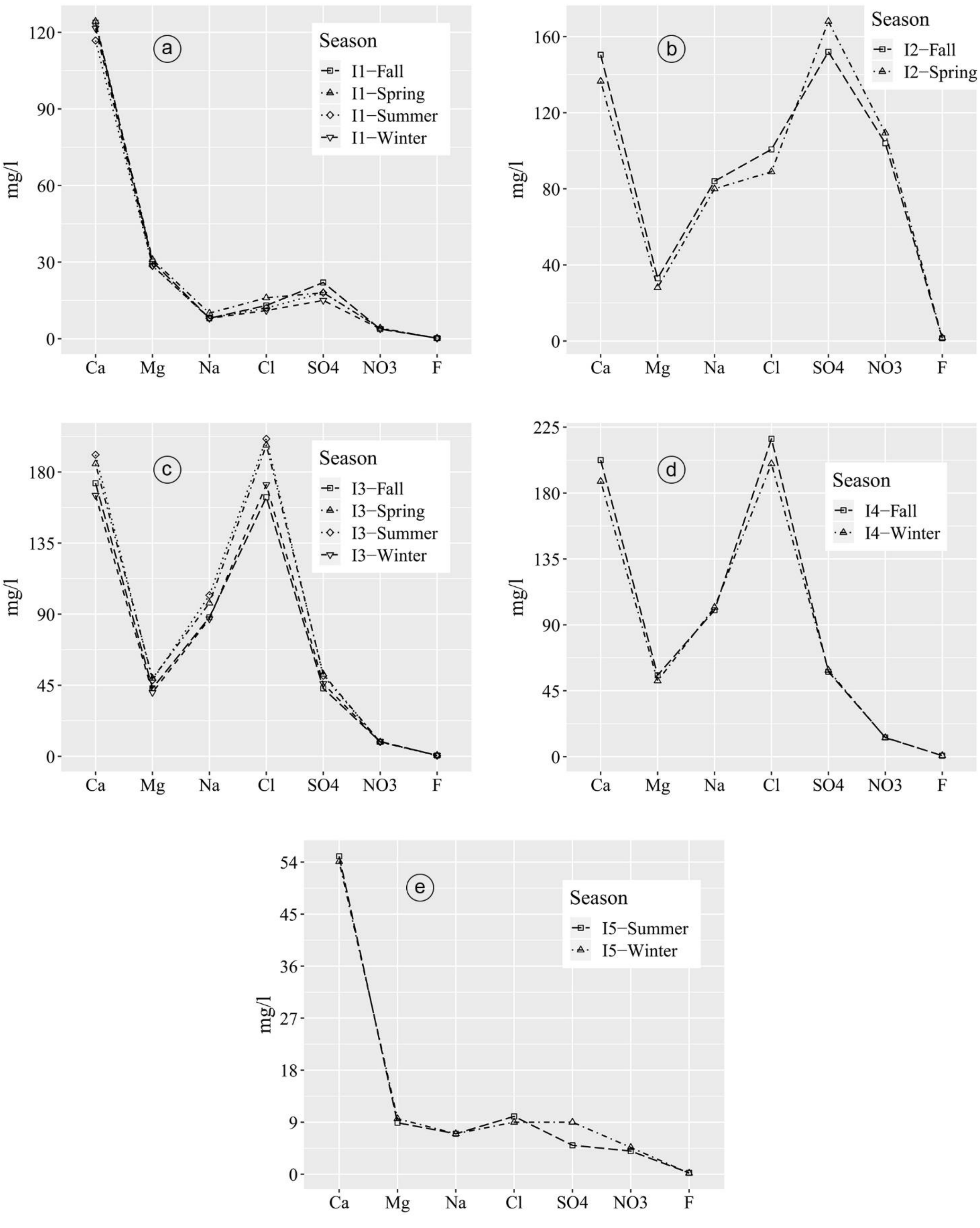


Fig. 6 Seasonal composition of the input waters in the water distribution system of Naples in 2017: (a) I1: ACO; (b) I2: AIR wells; (c) I3: Cancello well field; (d) I4: Campano Aqueduct and Ponte Tavano wells; and (e) I5: Serino Aqueduct.

Concentrations of major cations (i.e. Ca^{2+}, Mg^{2+}, Na^{+} and K^{+}) and anions (i.e. Cl^{-}, HCO_3^{-} and SO_4^{2-}) were plotted on trilinear diagrams (Fig. 7) to evaluate seasonal variation of tap water chemistry based on the dominant cations and anions. Figure 7a and b indicates that Ca^{2+} and HCO_3^{-} are the dominant cation and anion, respectively. In summer and fall, the relative concentrations of $Na^+ + K^+$ and Cl^- increase. Comparison of Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-}, Cl^-, NO_3^- and F^- measurements in the input waters (I1 to I5 in Fig. 6) and the tap waters (R1 to R7 and S1 to S3 in Fig. 8) help to interpret the seasonal changes in water chemistry. Except for S. Sebastiano (R2), ACO (S1) and Serino Aqueduct (S3), this seasonal trend is confirmed in the rest of tap waters (Fig. 8). Considerable Na^+ and Cl^- contents in the water from AIR wells (I2; Fig. 6b) might explain the high proportion of I2 in Capodimonte (R1) tap waters in summer and fall. This chemical signature can subsequently affect the relative abundance of sodium and chloride in the other waters due to the interconnections of the water reservoirs. Summer is the dry season in the study area with less precipitation and more groundwater exploitation because of high water demand. Therefore, the geochemical characteristics of pyroclastic materials in the Lufrano aquifer might significantly influence Na^+ and Cl^- contents in the water from AIR wells (I2) (Celico et al., 1994). Excess Na^+, K^+ and Cl^- in fall can be explained by the proportion of I2 in R1 along with the time lag between precipitation and groundwater recharge.

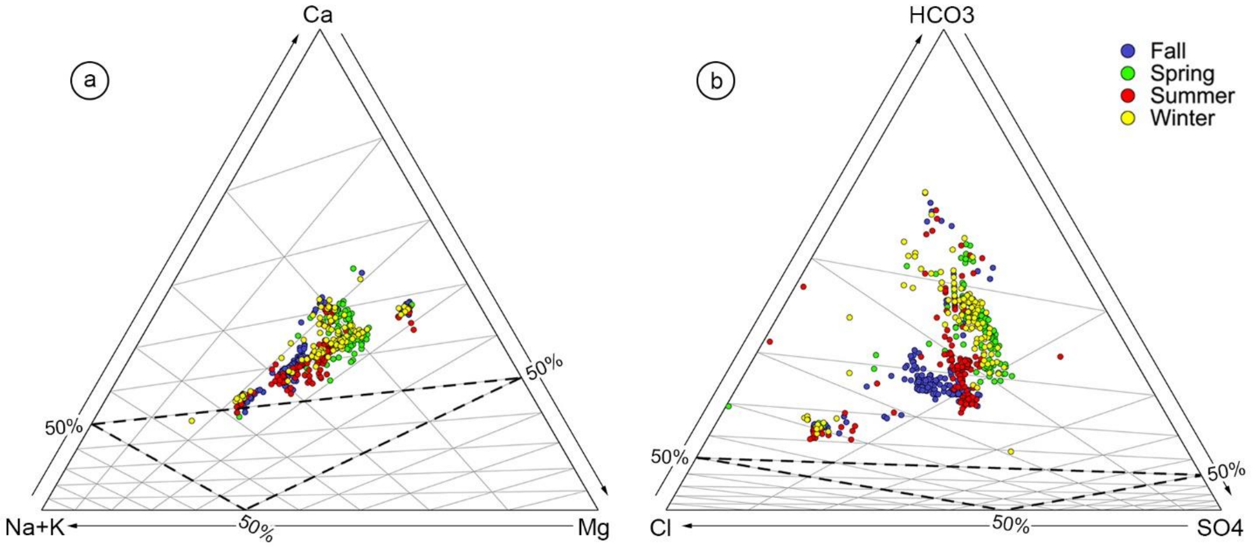


Fig. 7 Diagrams indicating seasonal composition of tap waters: (a) relative abundance of major cations; (b) relative abundance of major anions. The data are centered to get a better view of the pattern.

4.3. Construction materials of the reservoirs and tap water chemistry

Plotting the geometric mean of TDS against TH demonstrates that although the highest values are observed in the S. Sebastiano reservoir (R2) and Cancello well field, Campano Aqueduct and Ponte Tavano wells (S2), the lowest values are determined in the Serino Aqueduct (Fig. 9). The composition of the tap waters from Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli reservoirs (R3, R4, R5, R6 and R7, respectively) ranges between the chemical characteristics of the tap waters from the Serino Aqueduct and Capodimonte reservoir (Fig. 9). It suggests a proportional negative influence of the water from Capodimonte on the tap water quality because the reservoirs are extensively interconnected. In a decreasing order, Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli tap waters are feasibly influenced by Capodimonte water depending on the transfer rate of water from the Capodimonte reservoir to the Camaldoli reservoir (Fig. 9). Thus, there is more than one unknown factor restricting the ability to thoroughly evaluate the effect of building materials of the reservoirs on the water chemistry. Regardless of water storage in the S. Sebastiano reservoir, most of the R2 and S2 tap waters distributed in the eastern part of the city reveal elevated TH and TDS values (Fig. 10a and b). The urban water distribution network (Fig. 2) demonstrates that the tap waters originating from the S.

Sebastiano reservoir (R2) are occasionally mixed with the water from the Serino Aqueduct (I5) preventing a comprehensive interpretation of the influence of tuff on water quality. Figure 9 also shows that the collected tap waters are hard-fresh water and very hard-fresh water according to the TDS classification of Todd (1980) and the TH classification of Sawyer et al. (2003).

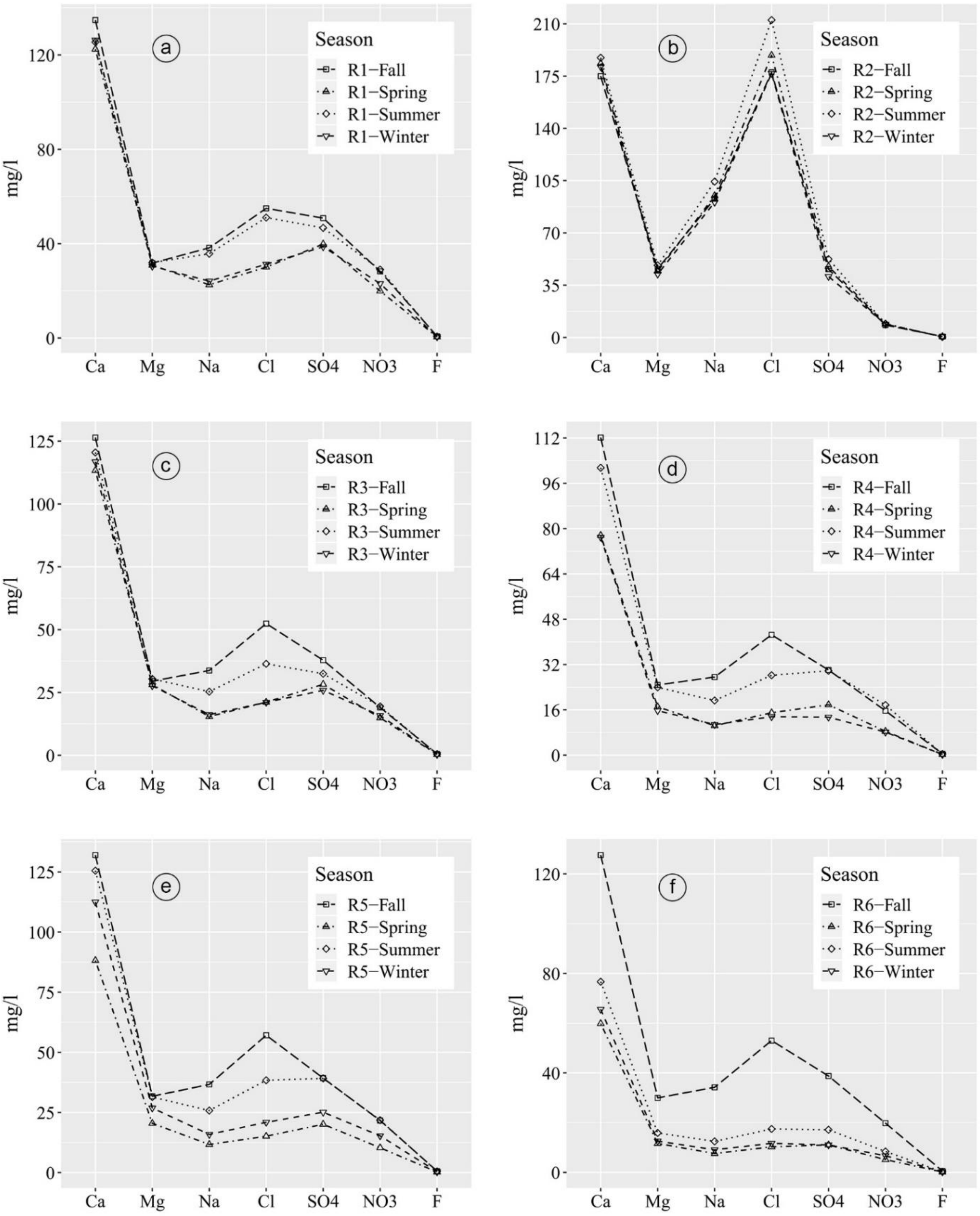


Fig. 8 Seasonal composition of the tap waters originated from: (a) R1: Capodimonte; (b) R2: S. Sebastiano; (c) R3: Scudillo; (d) R4: S. Giacomo; (e) R5: Pianura; (f) R6: Cangiani; (g) R7: Camaldoli; (h) S1: ACO; (i) S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; and (j) S3: Serino Aqueduct.

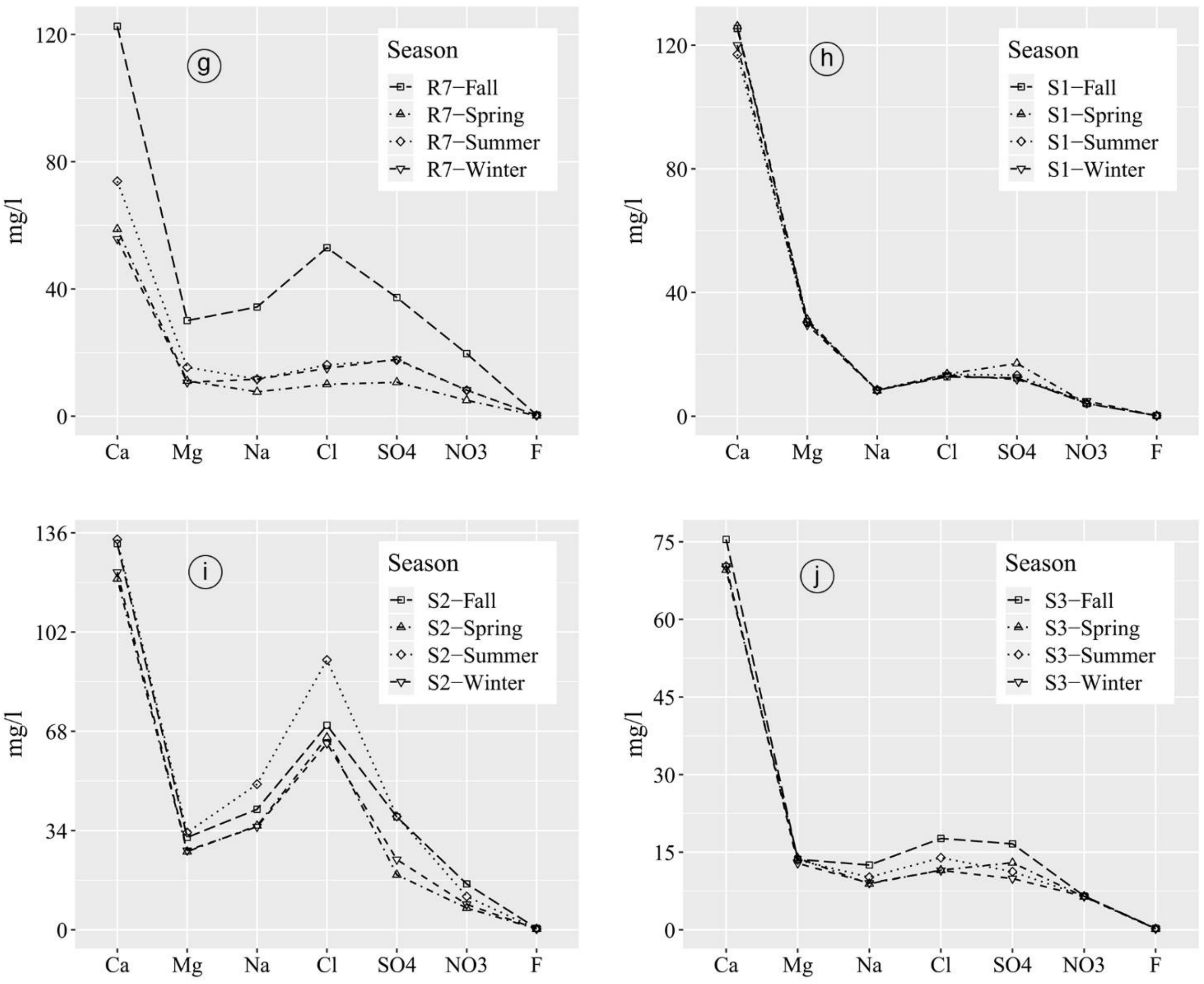


Fig. 8 (continued)

Because the Capodimonte reservoir does not receive water from many sources (Fig. 2), high values of anions, cations and the other parameters in R1 can be contributed by the input waters (I1 and I2) or the chemical interactions of the water with the concrete reservoir (Figs. 9 and 10). Compared with ACO (I1), elevated concentrations of the measured ions in AIR wells (I2) can be considered as the main source of chemicals in the Capodimonte tap waters (Fig. 6a and b). ACO (S1), which is not mixed with AIR wells (I2), has low contents of ions. A Portland cement plaster was applied to the interior tuffaceous walls of the Capodimonte reservoir. The hydrated phases of cement paste are not stable at pH values below 12 to 13. Prolonged contact of the porous material with near-neutral waters leads to dissolution and leaching of the soluble components (Alonso et al., 2006). The highly alkaline pore solution of a typical Portland cement paste leads to the beginning of the leaching process by removing Na^+ and K^+. It is followed by dissolution of portlandite and subsequent leaching of calcium from silicates (Gérard, 2000; Hidalgo et al., 2000; Taylor, 1997). Further, the dissolution/precipitation processes of aluminate phases including AFm (Al_2O_3–Fe_2O_3–mono: family of lamellar calcium alumina–ferrite hydrates), ettringite and calcite are affected (Faucon et al., 1997; Faucon et al., 1998). Na^+ and K^+ contents are higher in the water supplied by the Capodimonte reservoir (R1) than those of ACO (S1) probably suggesting the first stage of chemical reactions between the water and the concrete reservoir (Fig. 10k and l). It is in accordance with the indifferent concentrations of Ca^{2+} and Al in these two tap water sources (Fig. 10h and m). However, further investigation is needed to confirm this assumption. The reported leached depth of 5 to 10 mm in concretes submerged in still natural water for 100 years may explain the slow rate of the leaching process (Lagerblad, 1999). Thus, the elevated concentrations of $\mathrm{Cl^-}$, $\mathrm{SO_4^{2-}}$, $\mathrm{NO_3^-}$, and $\mathrm{F^-}$ in the water provided by Capodimonte highlight the unique role that AIR wells play in the water chemistry (Figs. 2, 8 and 10).

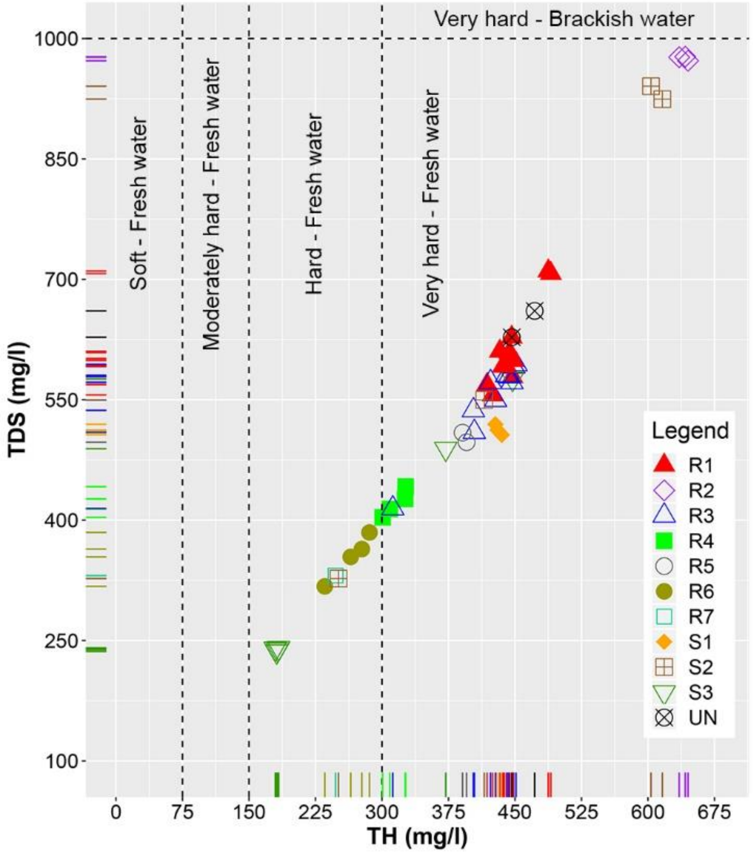


Fig. 9 Scatter plot of TDS vs. TH in the tap waters. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

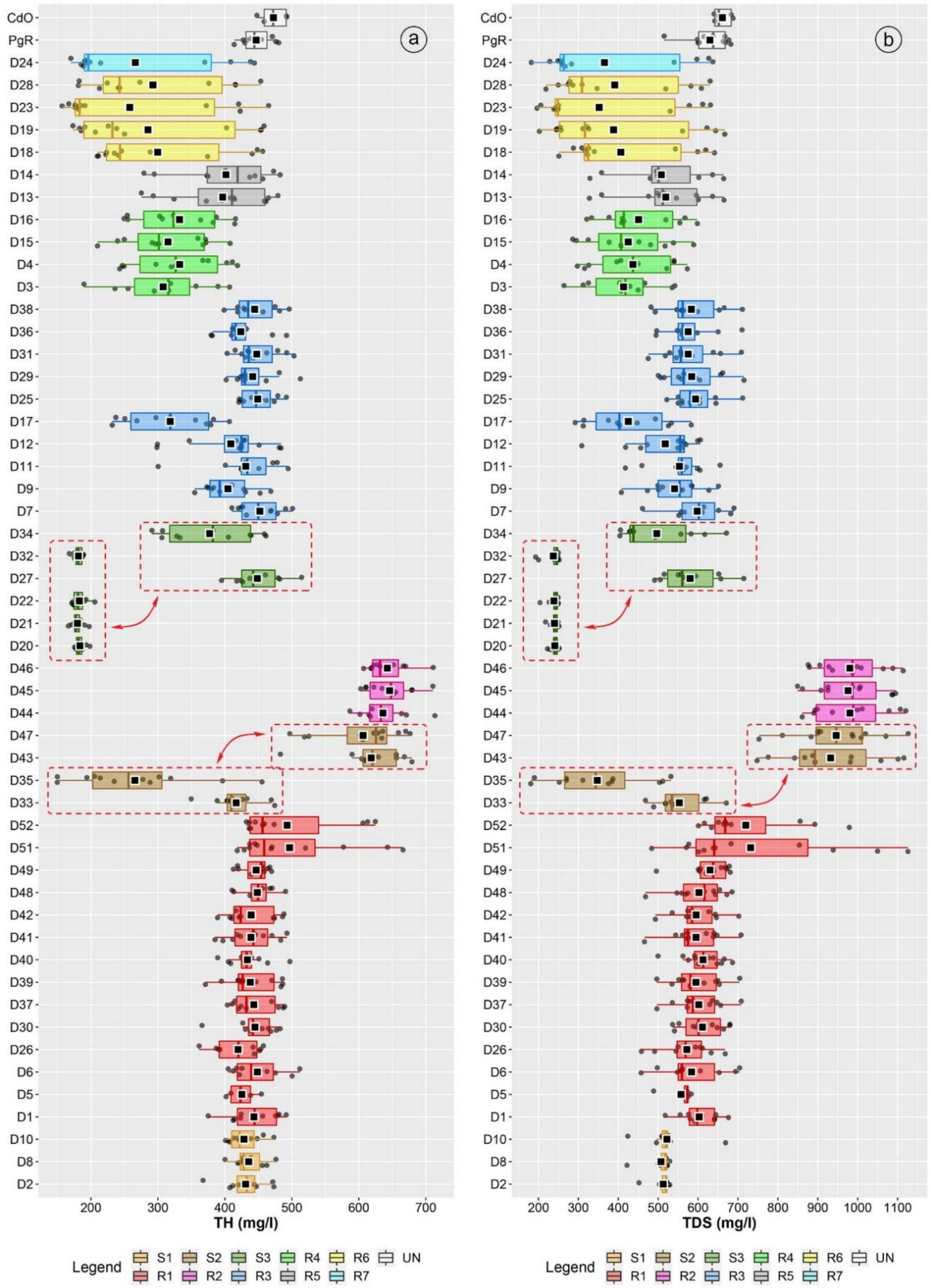


Fig. 10 Variation of (a) TH, (b) TDS, (c) As, (d) Ni, (e) V, (f) Fe, (g) Mn, (h) Al, (i) ClO_2^-, (j) Cl_2, (k) Na^+, (l) K^+, (m) Ca^{2+}, (n) Mg^{2+}, (o) HCO_3^-, (p) Cl^-, (q) SO_4^{2-}, (r) NO_3^-, (s) F^-, (t) EC, (u) turbidity, (v) pH and (w) temperature in the collected tap water samples. The gray circles indicate the monthly measurements and the black squares show the annual average values. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

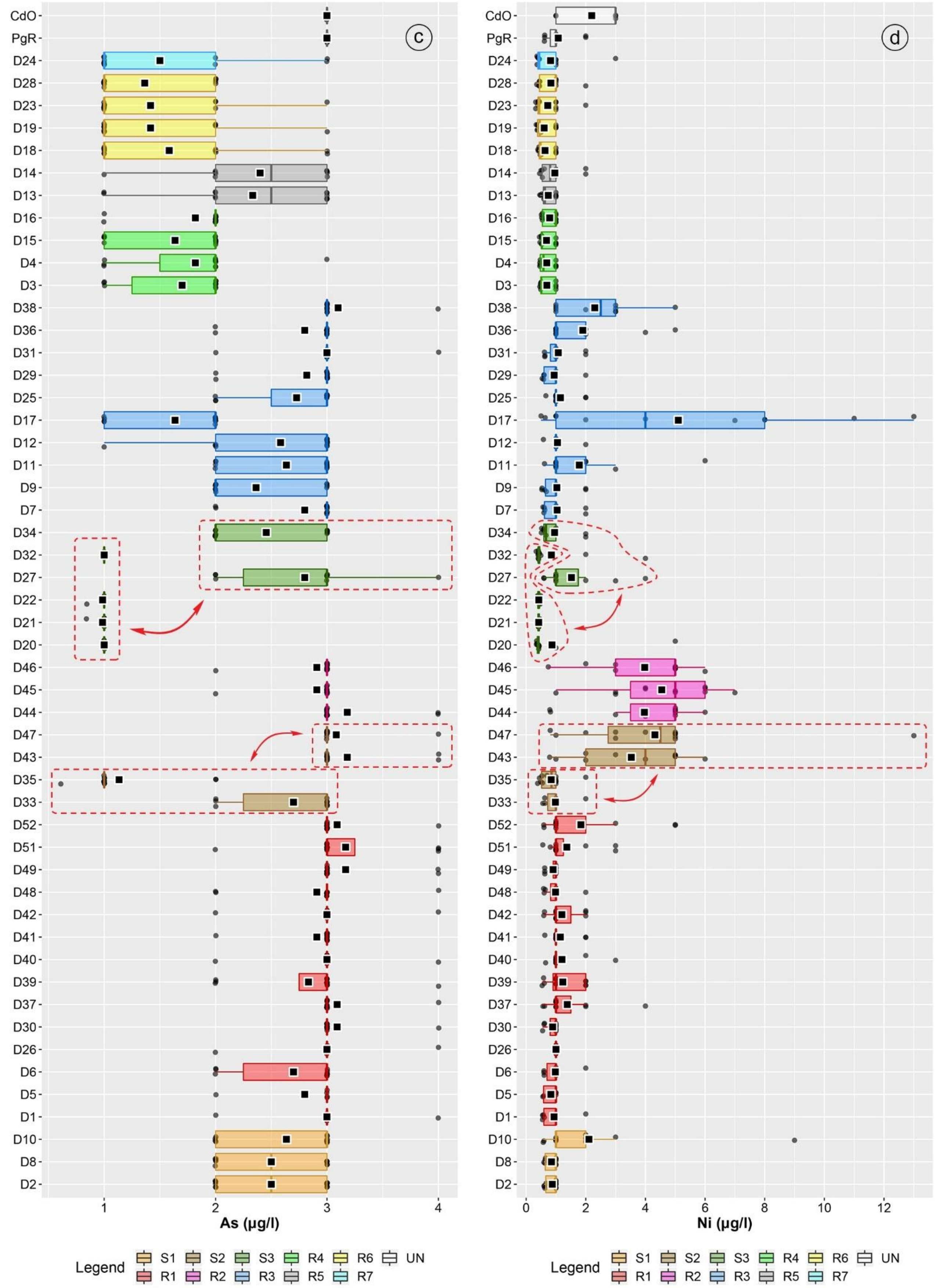


Fig. 10 (continued)

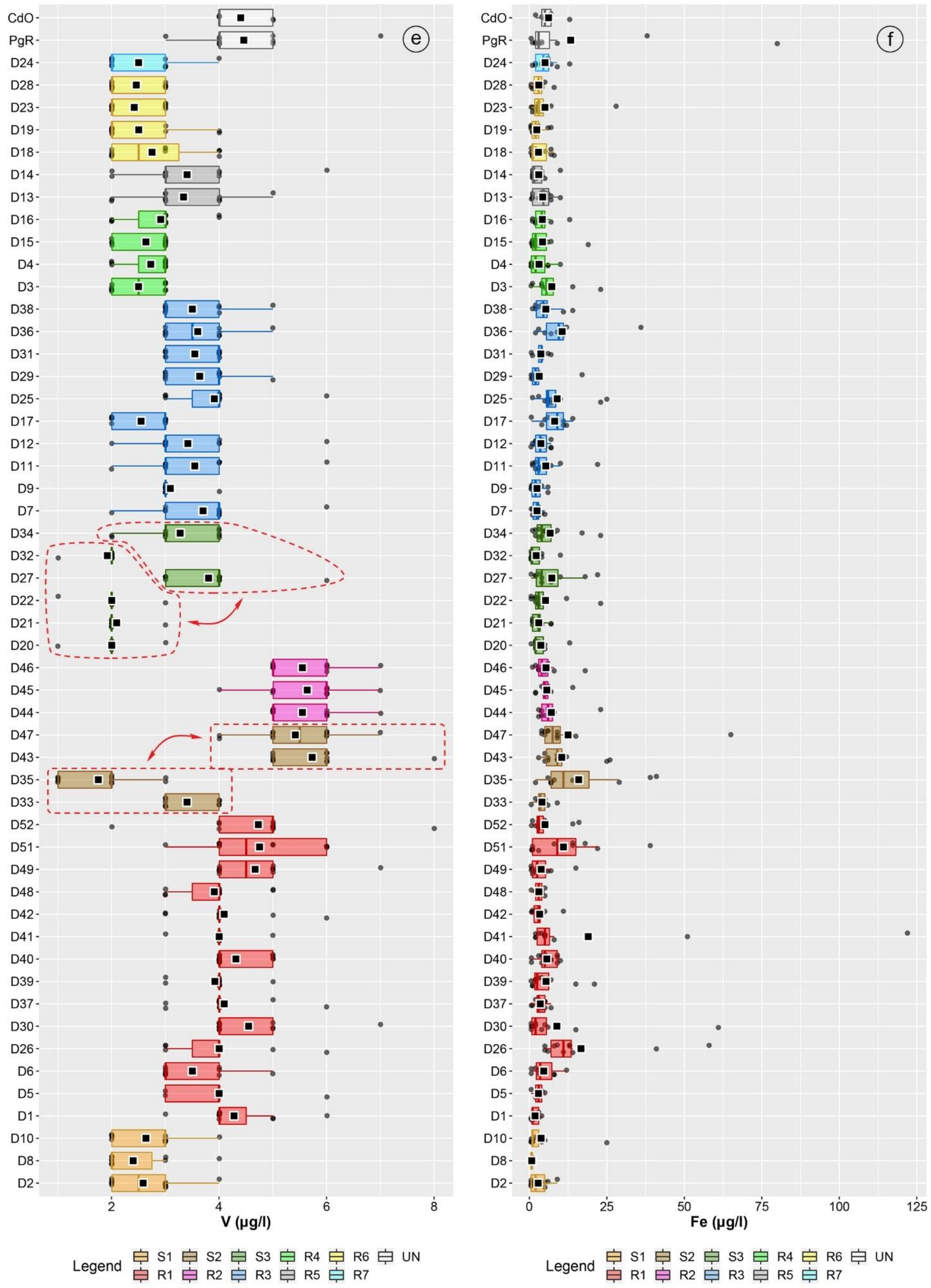


Fig. 10 (continued)

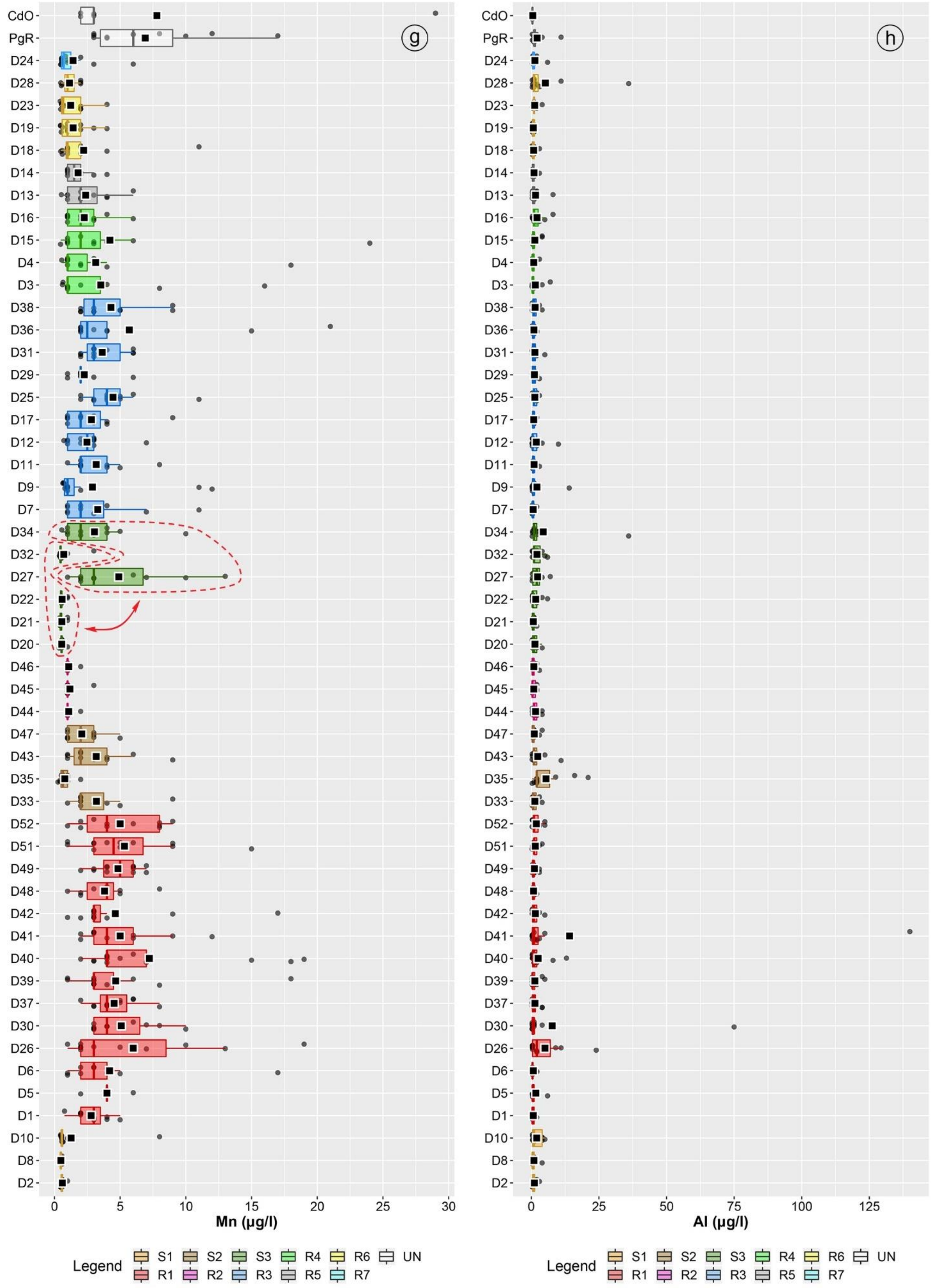


Fig. 10 (continued)

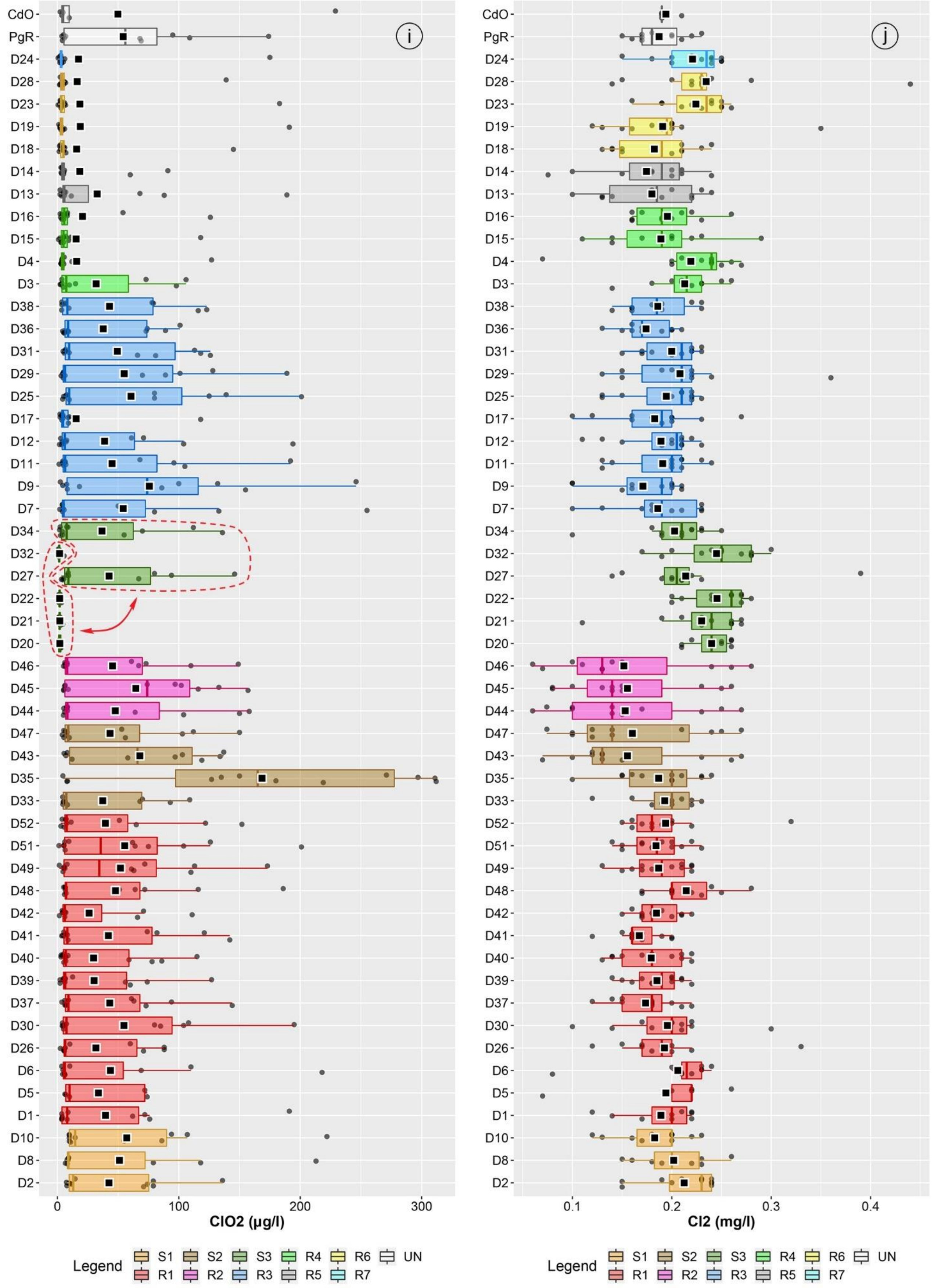


Fig. 10 (continued)

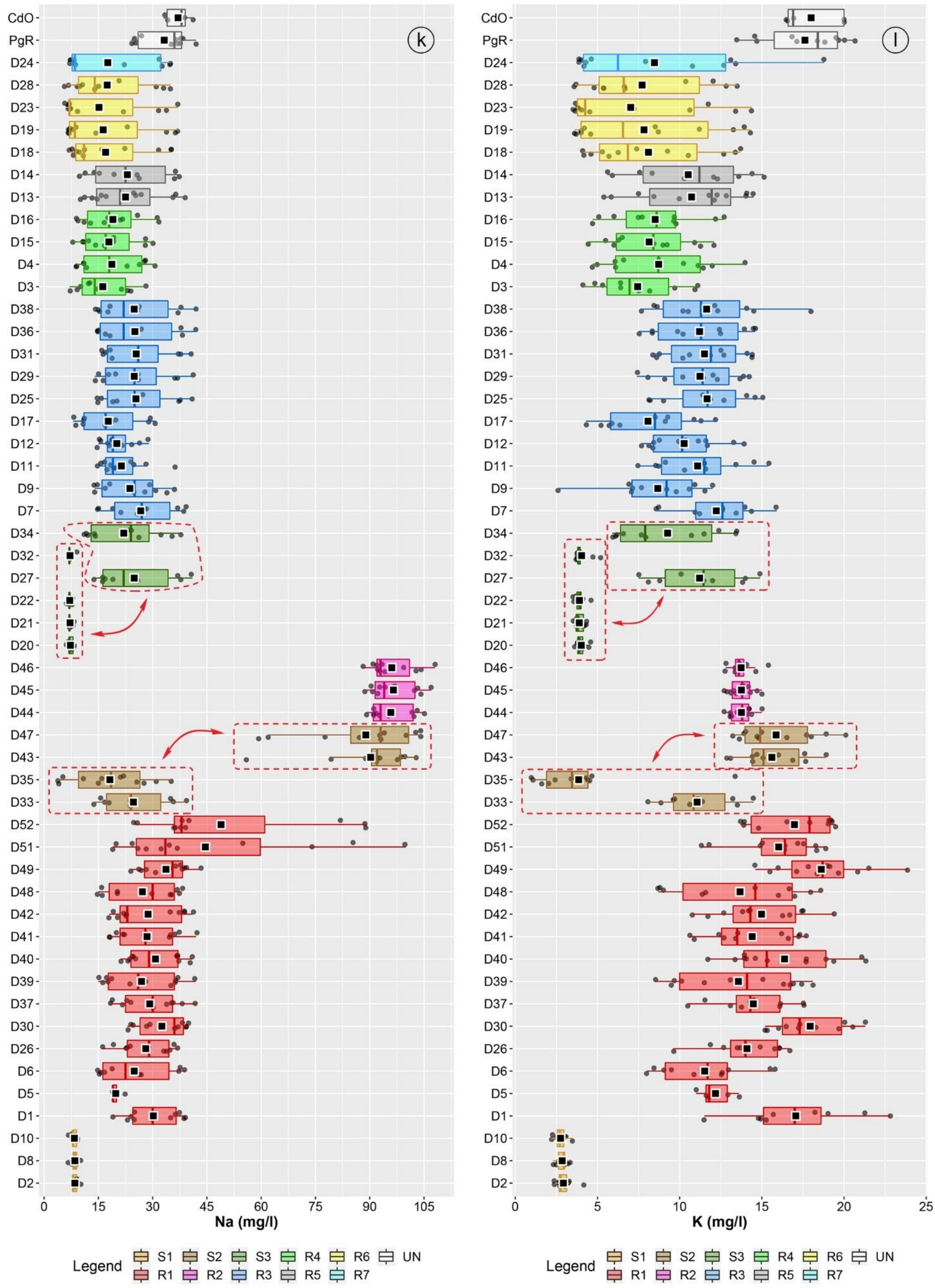


Fig. 10 (continued)

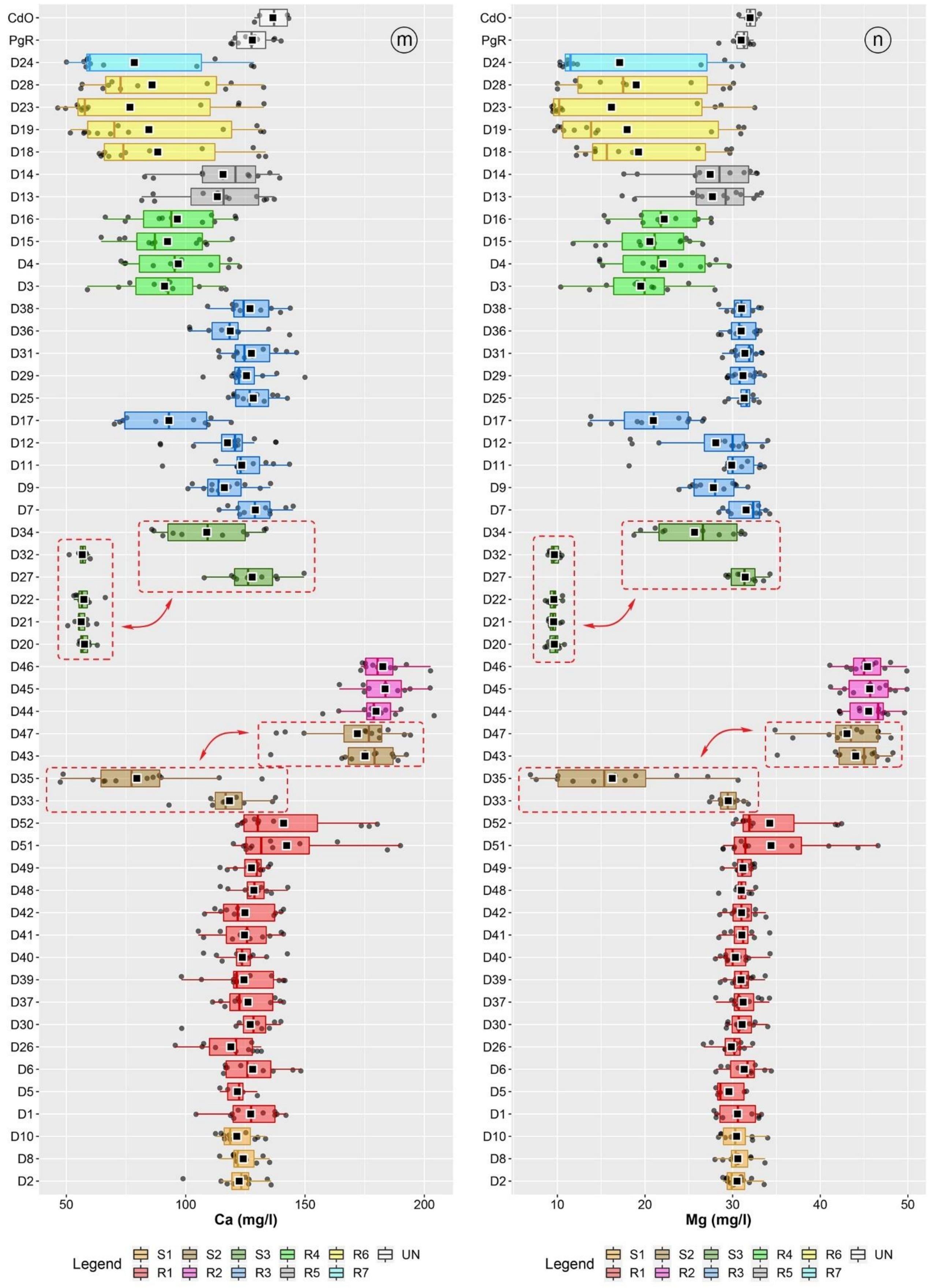


Fig. 10 (continued)

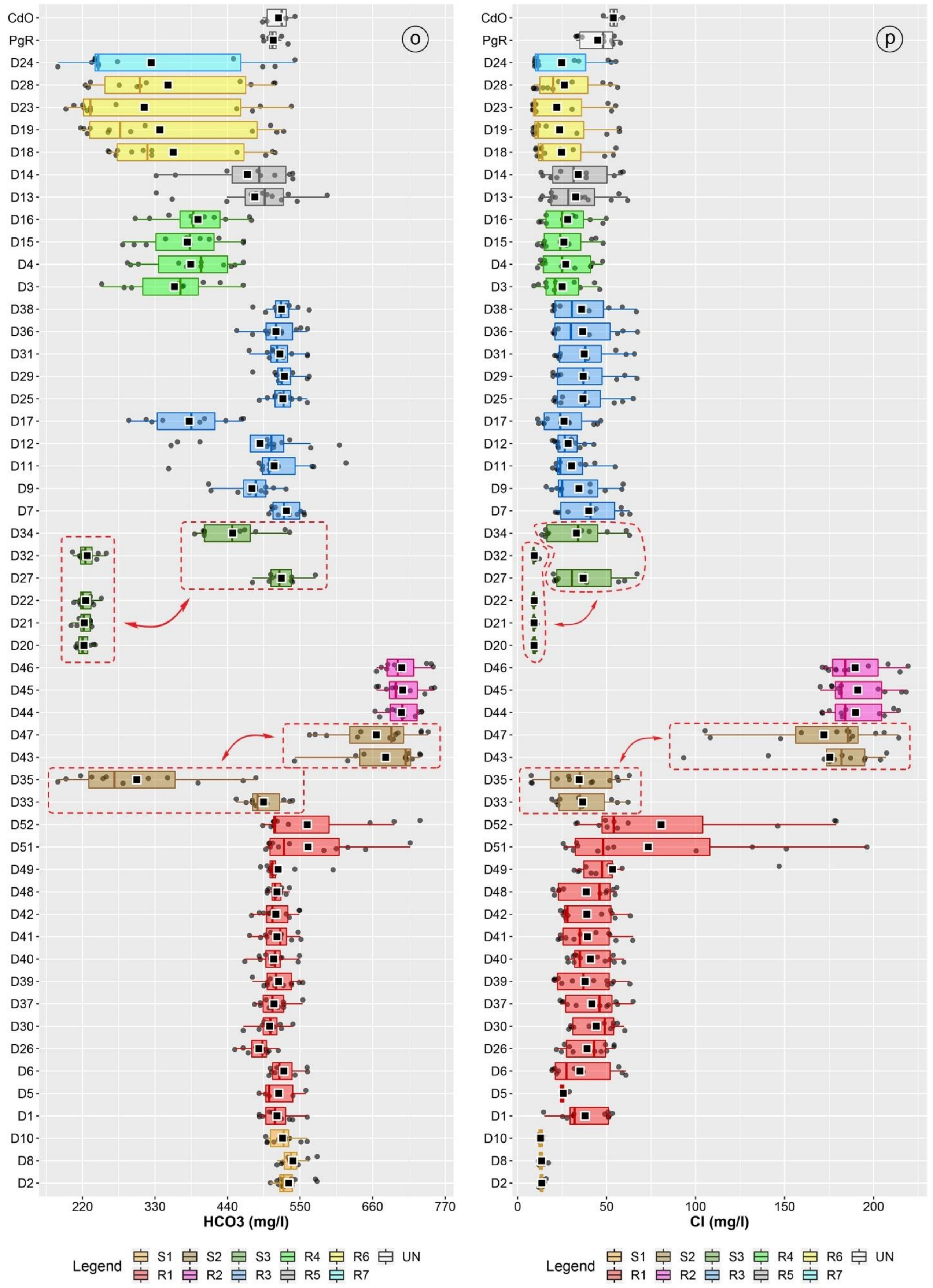


Fig. 10 (continued)

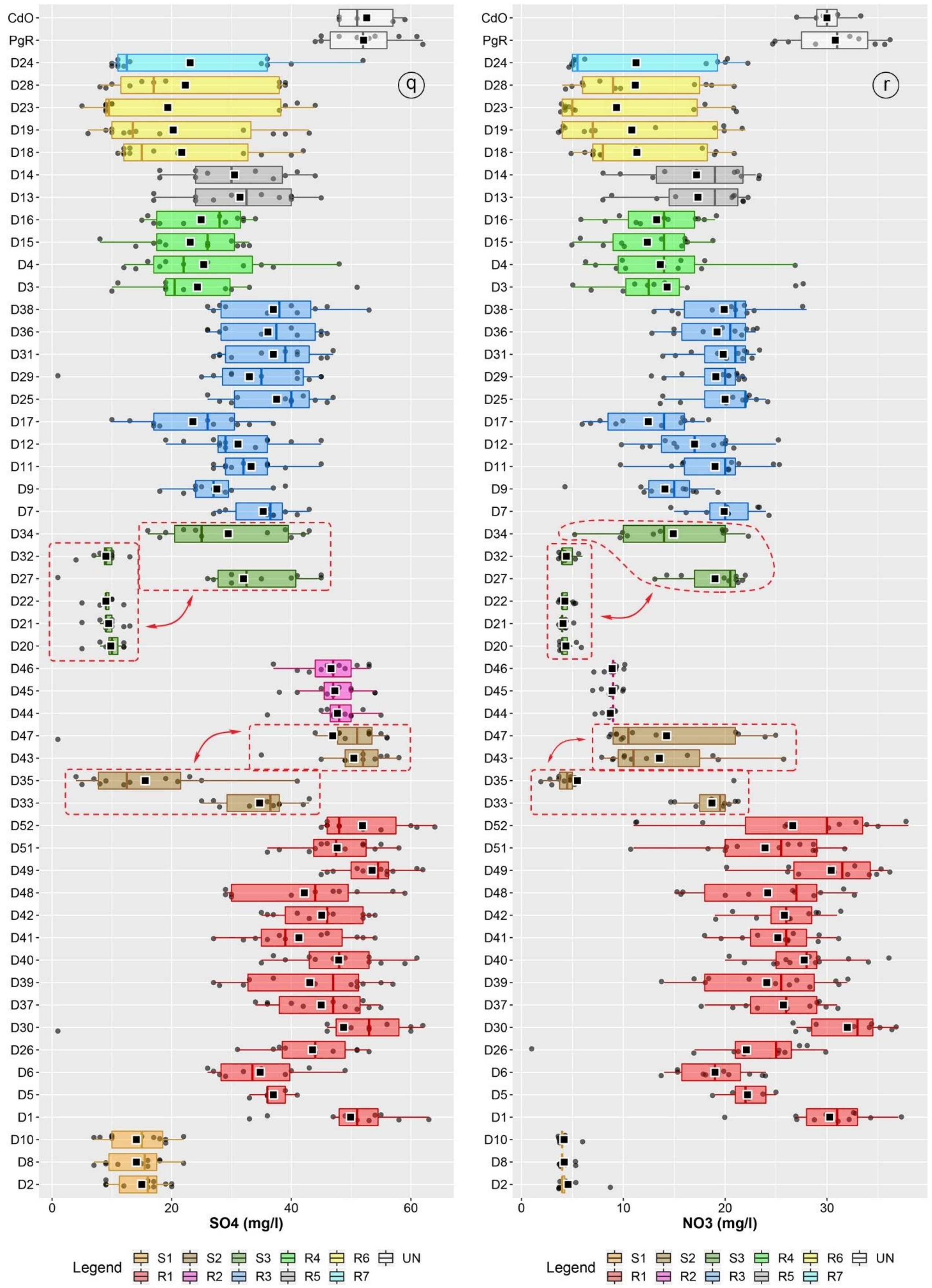


Fig. 10 (continued)

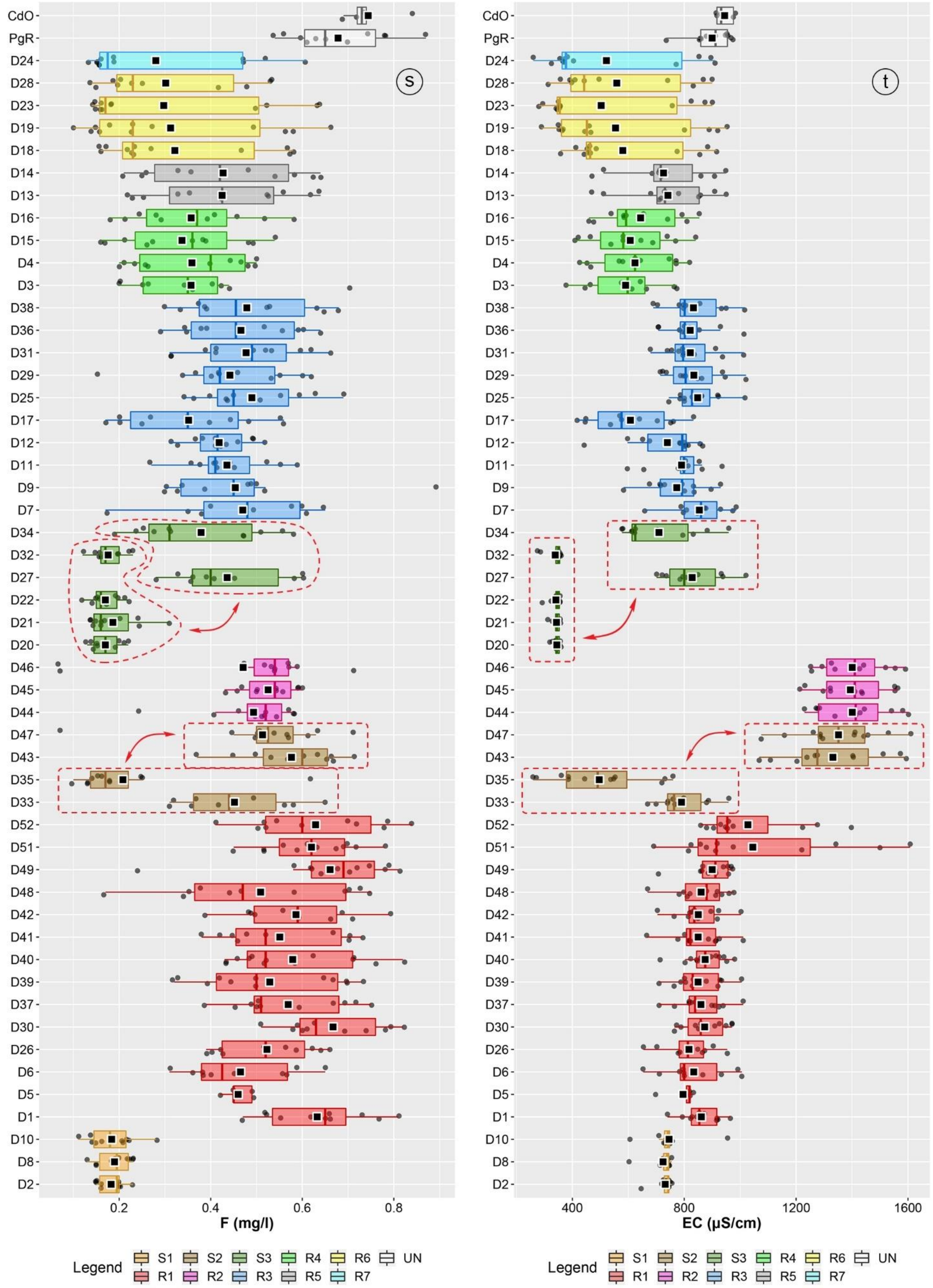


Fig. 10 (continued)

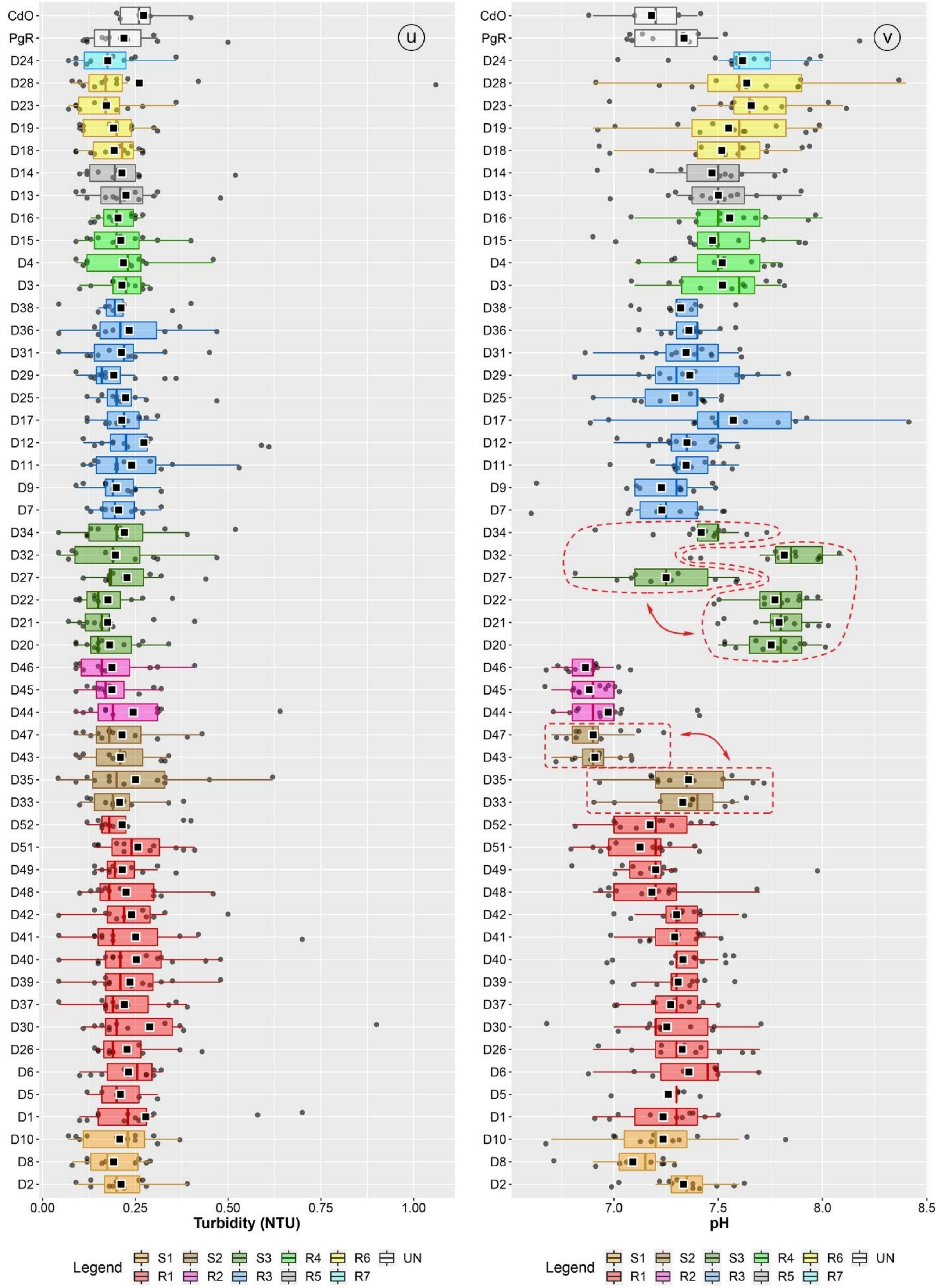


Fig. 10 (continued)

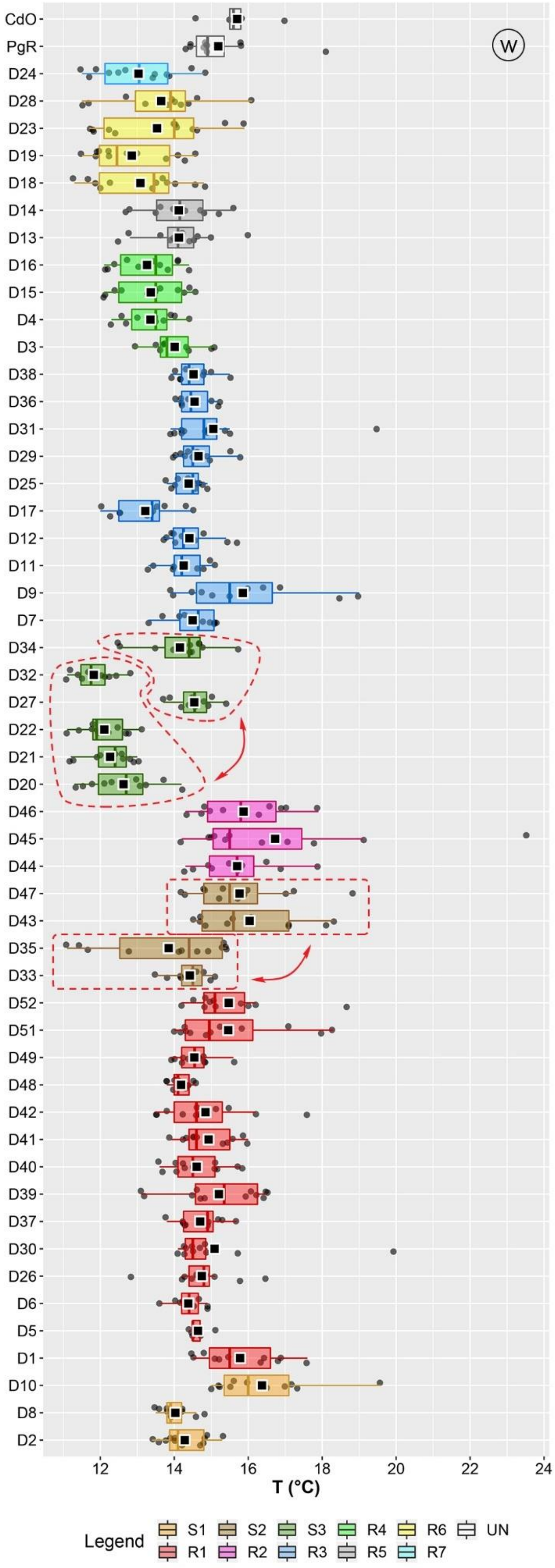


Fig. 10 (continued)

According to Fig. 10a and b, TH and TDS values of D33 and D35 are strikingly different from those of D43 and D47. Likewise, variation of TH and TDS in D27 and D34 is completely different from D20, D21, D22 and D32 supplied by the same source (Fig. 2). Similar surprising patterns in the other parameters (Fig. 10c to w) might indicate lack of information about the network structure in some parts of Naples and stress the need for characterizing the source of tap water through the existing dataset.

4.4. Chemical fingerprint of the source in the water distribution network

A clr-biplot was used (Fig. 11) to investigate the compositional variability along with the associations between the parts of a (sub)composition and to identify the probable natural sources of the water provided for human consumption. The first two principal components explain 98% (62 and 36%, respectively) of the total cumulative variability being an acceptable approximation for the 5-dimensional scatterplot of the clr-transformed data. Chloride and nitrate indicate the greatest compositional variability; however, sulfate shows the smallest variability. Ca^{2+} and HCO_3^- are significantly correlated, but their correlation with SO_4^{2-} is negligible. The clr-biplot indicates that the tap water samples can be discriminated by three groups of variables. The samples from the S. Sebastiano reservoir (R2) together with the Cancello well field, Campano Aqueduct and Ponte Tavano wells (S2) are plotted around the “clr.Cl” ray introducing chloride as the main chemical fingerprint of the waters. Sulfate and nitrate characterize the hydrochemistry of Capodimonte water (R1) whilst calcium and bicarbonate are the main characteristics of S1 and S3 waters (ACO and Serino Aqueduct, respectively) (Fig. 11). Occurrence of the other samples (Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli) between these two endmembers confirms the results in section 4.3 and Fig. 9.

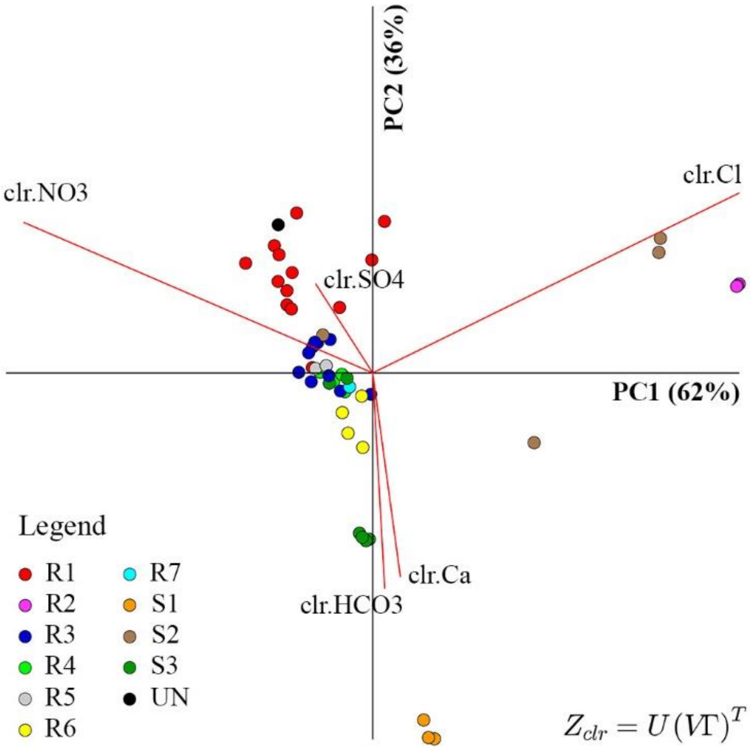


Fig. 11 The clr-biplot of the selected ions retaining 98% of the total variability. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

Visualizing the spatial distribution of the factor scores of the principal components, after appropriate classification, also depicts the chemical endmembers (Fig. 12). Cl^- and NO_3^- are the ions with significant factor loadings in the first principal component (PC1) (0.718 and −0.684, respectively). According to Fig. 12a, the highest PC1 factor scores in the eastern part of the city correspond to high chloride content in R2 and S2 tap waters. Nonetheless, nitrate enrichment in most of the water samples from Capodimonte, Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli is reflected in their lowest factor scores. In terms of Cl^- and NO_3^-, the S1 and S3 tap waters (Serino Aqueduct and ACO, respectively) lie between the abovementioned endmembers (Fig. 12a).

On the other hand, the second principal component (PC2) accounts for significant negative factor loadings of Ca^{2+} and HCO_3^- (−0.525 and −0.555, respectively), and a high positive factor loading of Cl^- (0.463). The group with the highest factor scores is representative of the Cl-bearing R1, R2 and S2 tap waters. However, the Serino Aqueduct (S3) and ACO (S1) are characterized with the lowest factor scores due to the unique clr-transformed calcium and bicarbonate values. The rest of the samples feasibly depict a mixed source (Fig. 12b).

4.5. Hierarchical cluster analysis and tap-water supply map

Hierarchical cluster analysis of the tap water samples collected in Naples was performed and the results are presented in Fig. 13. The seasonal differences in the groups suggest that some urban water reservoirs cannot be considered in isolation as single water bodies. The water samples from Capodimonte (R1) and ACO (S1) are clustered into distinct categories indicating the significant role of the water from AIR wells (I5) in the composition of water from the Capodimonte reservoir. In comparison with the other seasons, the samples from Capodimonte and Scudillo (R1 and R3, respectively) are subgroups of the same cluster in spring indicating that Capodimonte makes the greatest contribution to the amount of water in the Scudillo reservoir in this season (Fig. 13b).

Depending on the seasonal interconnections between the urban water reservoirs, the water provided by S. Giacomo (R4), Pianura (R5), Cangiani (R6) and Camaldoli (R7) can be chemically similar to the water from the Serino Aqueduct or Scudillo (Fig. 13). Considering cluster D of HCA in spring (Fig. 13b), the Serino Aqueduct is the main source of the tap water supplied by the abovementioned reservoirs (S. Giacomo, Pianura, Cangiani and Camaldoli). It seems that the volume of water transferred from Scudillo to S. Giacomo (Fig. 2) gradually increases in winter, summer and fall (Fig. 13). Pianura is the first reservoir that receives Scudillo water in winter (all in cluster B) when there is a low rate of pumping water from Scudillo to S. Giacomo (Fig. 13a). By raising water input from Scudillo in summer, the waters from S. Giacomo, Pianura and Scudillo are grouped in the same cluster owing to their similar chemical characteristics (Fig. 13c). S. Giacomo, Pianura, Cangiani and Camaldoli tap waters (R4, R5, R6 and R7, respectively) are identical with Scudillo tap water (R3) in fall when the water contribution from Scudillo is maximum (Fig. 13d). The abovementioned discussion might suggest that the Scudillo reservoir is connected to the Pianura reservoir because it is the first one affected by Scudillo water in winter and it occurred in the cluster of Scudillo in three seasons.

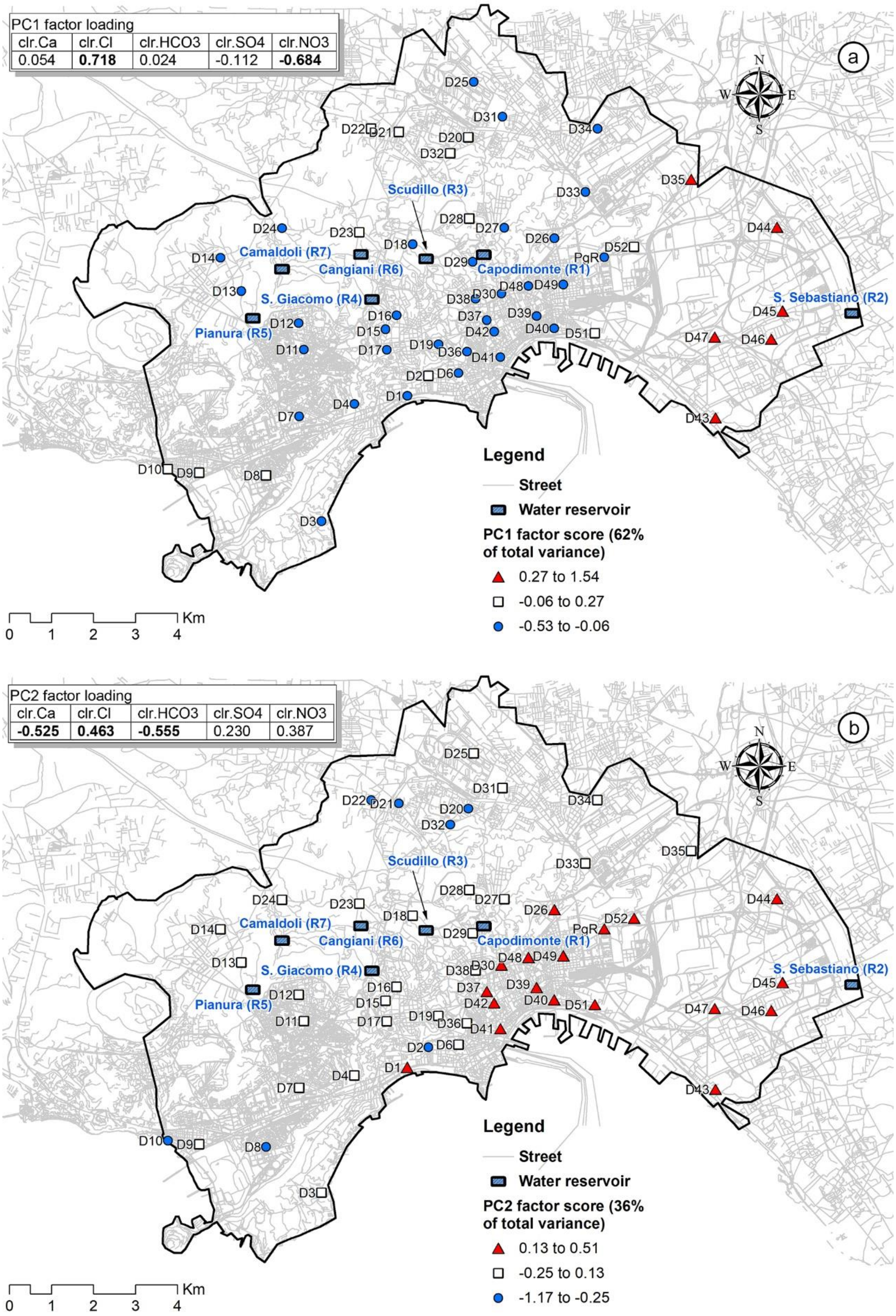


Fig. 12 Factor score maps for the first two principal components explaining 98% of the total variance in the dataset: (a) the first principal component (PC1); and (b) the second principal component (PC2). In order to find an interpretable classification, the upper and lower values of each class were adjusted to match the spatial distribution of different tap waters with the tap-water supply scheme (Fig. 2).

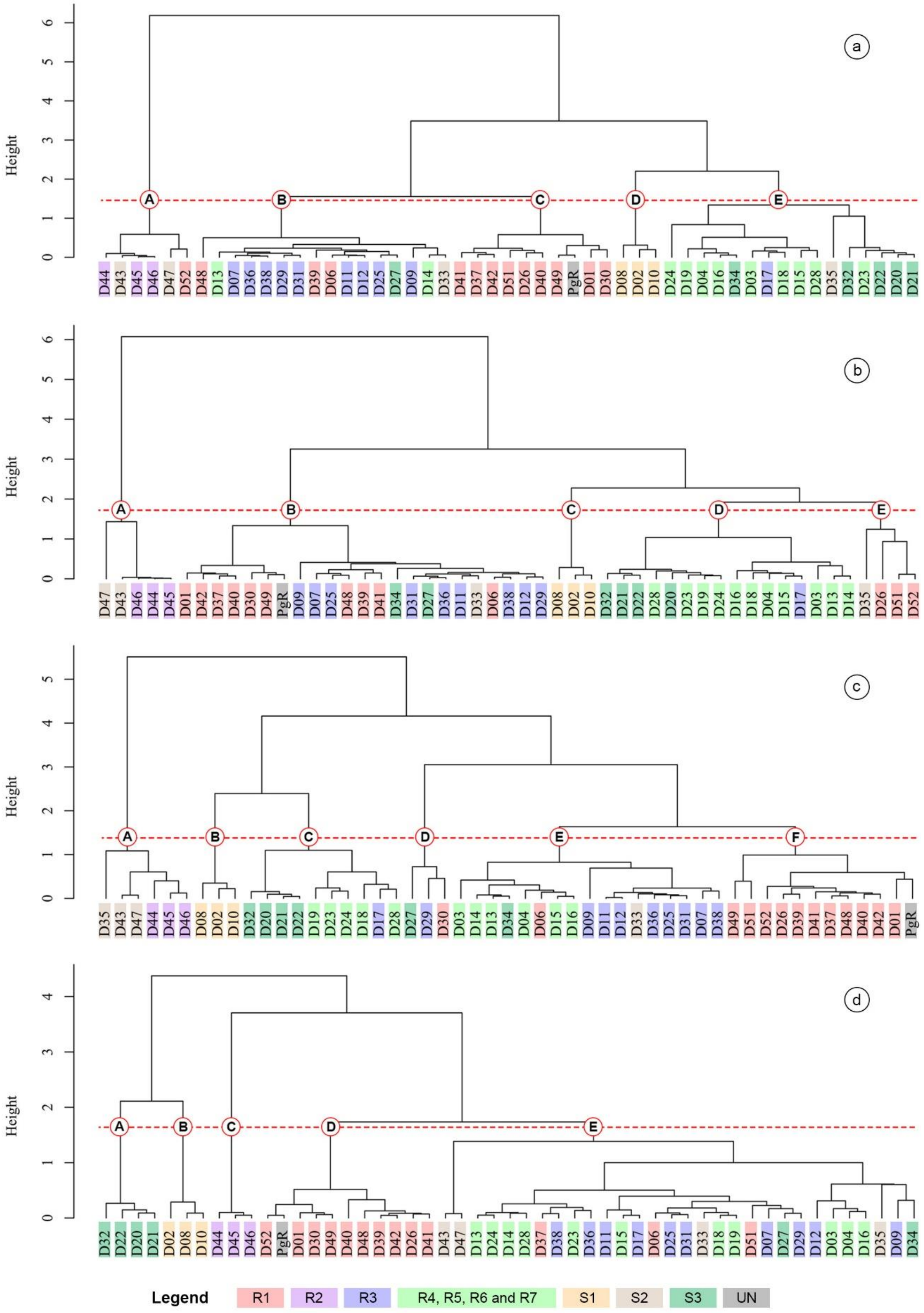


Fig. 13 Hierarchical cluster analysis of the tap water samples using the clr-transformed data: (a) winter; (b) spring; (c) summer; and (d) fall. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

Whenever excess water from the Serino Aqueduct (I5) is available, it is mixed with the water in the S. Sebastiano reservoir (R2) to improve the water quality (Fig. 2). In summer, S. Sebastiano probably receives no water from the Serino Aqueduct and consequently most of the R2 and S2 tap water samples are clustered together. However, D35 (together with D43 and D47 in fall) is clustered with group E in winter, spring and fall (Fig. 13). It seems that the pipeline transferring the excess water from the Serino Aqueduct to the S. Sebastiano reservoir (Fig. 2) is not directly connected to the reservoir since transformed values of the ions and physicochemical parameters in D35 (in all seasons except for summer) along with D43 and D47 (in fall) are between the corresponding values of the S. Sebastiano reservoir and Serino Aqueduct (Fig. 14). The non-transformed values of the parameters are also some pieces of supporting evidence (Fig. 10).

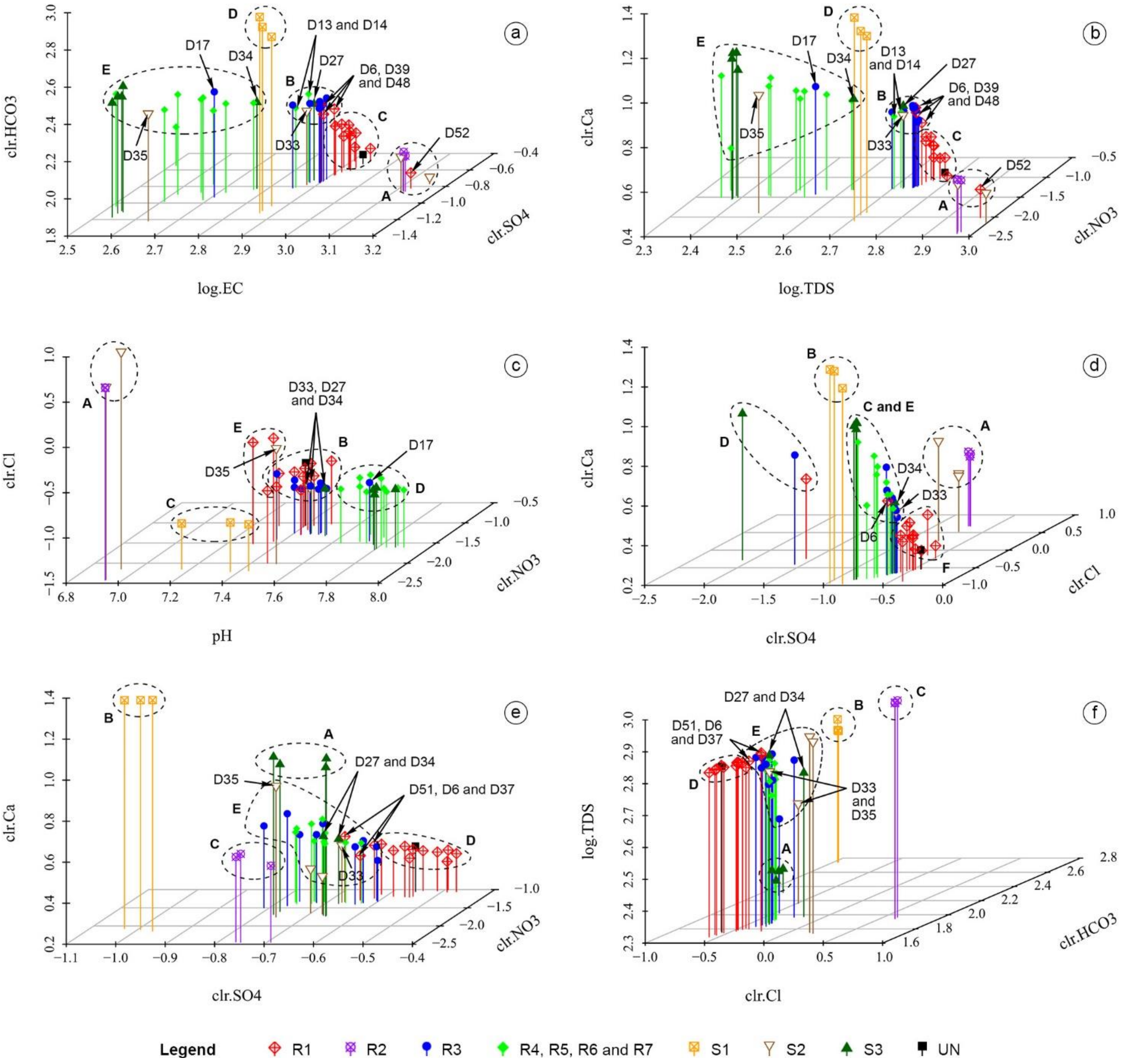


Fig. 14 3D scatter plots of the selected variables: (a and b) winter; (c) spring; (d) summer; and (e-f) fall. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme. The capital letters (A to E) represent the clusters in Fig. 13.

Except for cluster B in spring where the samples from Capodimonte (R1) and Scudillo (R3) are agglomerated, D6 is never placed in a group with the majority of samples from Capodimonte (Fig. 13). D33 is also clustered with Scudillo instead of R2 and S2 in different seasons. In addition, D27 and D34 do not occur in the group of the Serino Aqueduct (S3). It is consequently proposed that the source of tap water in D6, D27 and D33 is the Scudillo reservoir; however, D34 is representative of S. Giacomo (including Pianura, Cangiani and Camaldoli). On the other hand, D17 which is supposed to receive water from Scudillo is statistically similar to S. Giacomo, Pianura, Cangiani and Camaldoli. It is in accordance with the boxplots (Fig. 10) and 3D scatterplots (Fig. 14) depicting the chemistry of the statistically analyzed variables and samples.

The chemical fingerprint of the PgR sample which does not exist in the tap-water supply scheme is analogous to the water from the Capodimonte reservoir (Figs. 10, 13 and 14). The similarity of Ca^{2+}, HCO_3^-, SO_4^{2-}, NO_3^-, EC and TDS transformed values in D35 with those of the Serino Aqueduct, S. Giacomo, Cangiani and Camaldoli led to the occurrence of D35 in group E in winter. Nevertheless, in spring, D35 is clustered with D26, D51 and D52 because of almost the same Cl^- and pH values. The presence of D27, D29 and D30 in group D in summer can be verified by their roughly equal "clr.Cl" and "clr.SO4" (Fig. 14). Considering the transformed values of Ca^{2+}, HCO_3^-, Cl^-, NO_3^- and TDS in fall, it makes sense that D43 and D47 occur in cluster E with the tap waters from the Serino Aqueduct, S. Giacomo, Pianura, Cangiani and Camaldoli (Figs. 13 and 14).

The information in Fig. 13 is spatially presented by considering some points (Fig. 15): (1) the anovementioned interpretation for HCA is taken into account to avoid the unreliable information in the water supply scheme; (2) with the exception of group E in spring and group D in summer, the name given to each cluster follows the majority of the samples in the group; (3) whenever tap water samples from the Serino Aqueduct occur with S. Giacomo, Pianura, Cangiani and Camaldoli in the same cluster, the cluster is called the Serino Aqueduct which is the main water source; and (4) the group including water samples from Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli is called Scudillo because of water transfer from Scudillo to the aforementioned reservoirs. It is evident that depending on water transport from one reservoir to another, seasonal maps for water distribution should be considered (Fig. 15). However, all the four maps are combined to generate a unique map showing the city areas with a distinct source of tap water (Fig. 16). City areas are identified by the samples with the same HCA result in at least three seasons. In the other parts without zonation, seasonal changes should be considered for better water management. Boxplots of the monthly measurements (Fig. 10) and spatial representation of the seasonal HCA (Figs. 15 and 16) reveal that the water with the least content of cations and anions (i.e. the water from the Serino Aqueduct) is distributed in the northern parts of Naples. Nevertheless, the residents living in east consume the water with the highest content of dissolved chemicals. The extent of the S. Sebastiano area (R2 in Fig. 16) is almost equivalent to the city area considered by Giugni et al. (2008) for hydraulic simulation of the "Napoli Est" water distribution system. In the DMA design, the WDS is usually divided into more homogenous parts with integration of visual examination and technical experience. However, it is complicated in large WDSs with numerous user nodes and pipes (Scibetta et al., 2014). It might explain that the proposed approach can provide some ideas for engineers and managers to divide the complex pipe networks into unique areas prior to hydraulic modelling. Considering the high maintenance operations and the noticeable water loss in the eastern sector of the city (Giugni et al., 2008), it is highly important to find a solution and manage the WDS in the study area because the problem is feasibly more significant in the historic center of Naples. If it is possible to identify the system's properties (e.g., the network structure and the interactions between its components) in at least some areas defined in the current investigation (Fig. 16), setting up some modelling techniques with minimum prerequisites can further clarify the efficiency of the proposed methodology.

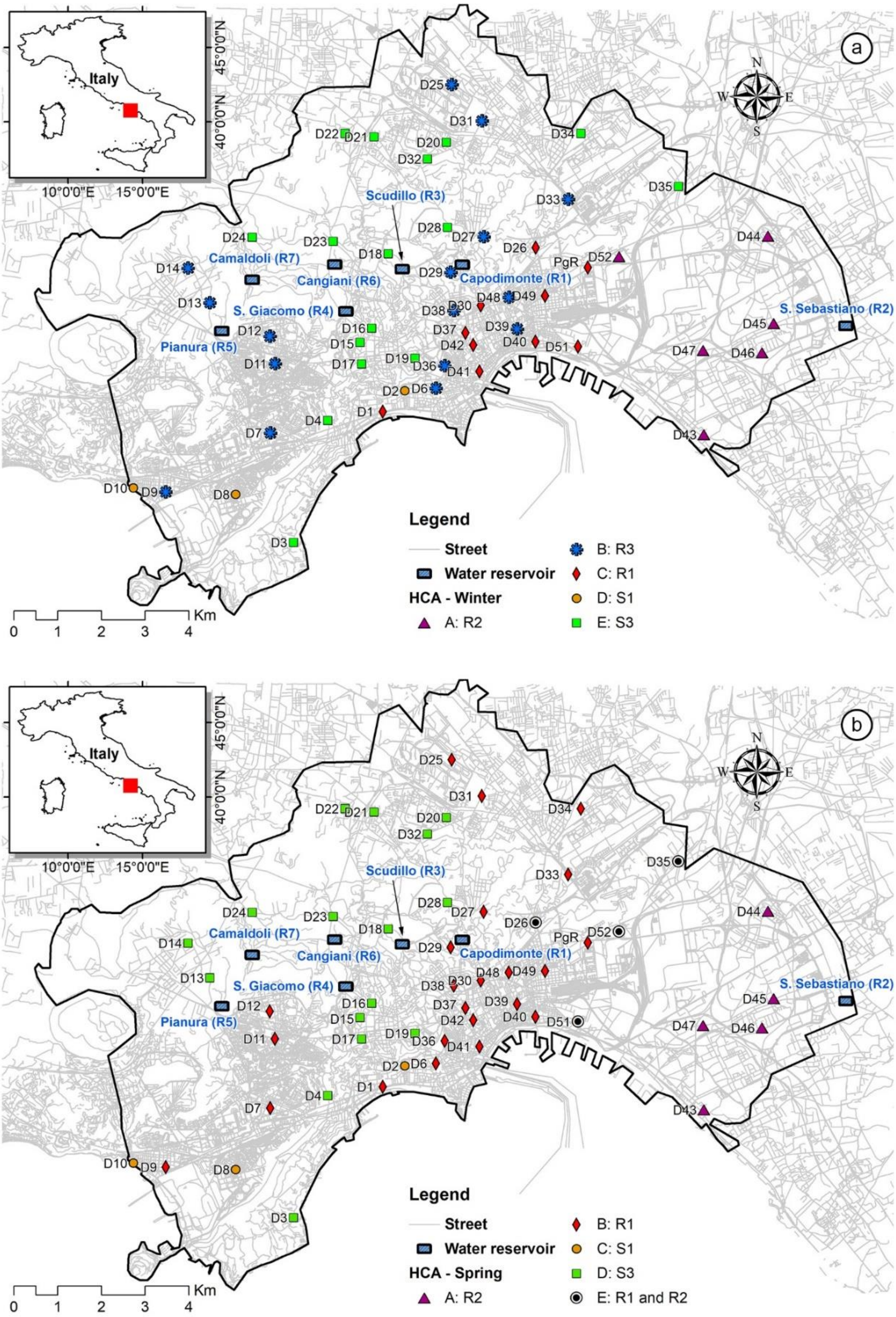


Fig. 15 Spatial representation of the results of seasonal hierarchical cluster analysis: (a) winter; (b) spring; (c) summer; and (d) fall. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; S1: ACO; and S3: Serino Aqueduct.

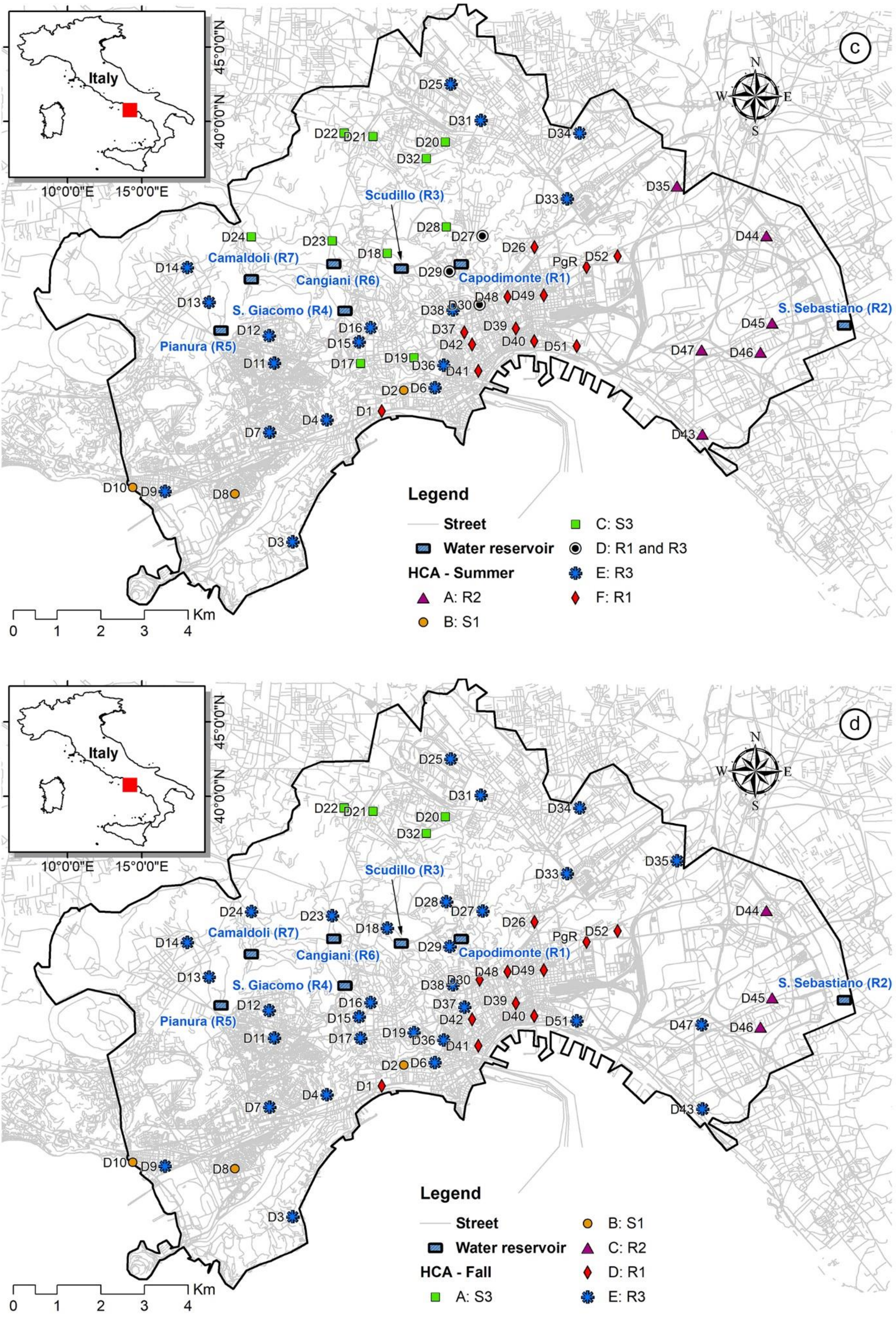


Fig. 15 (continued)

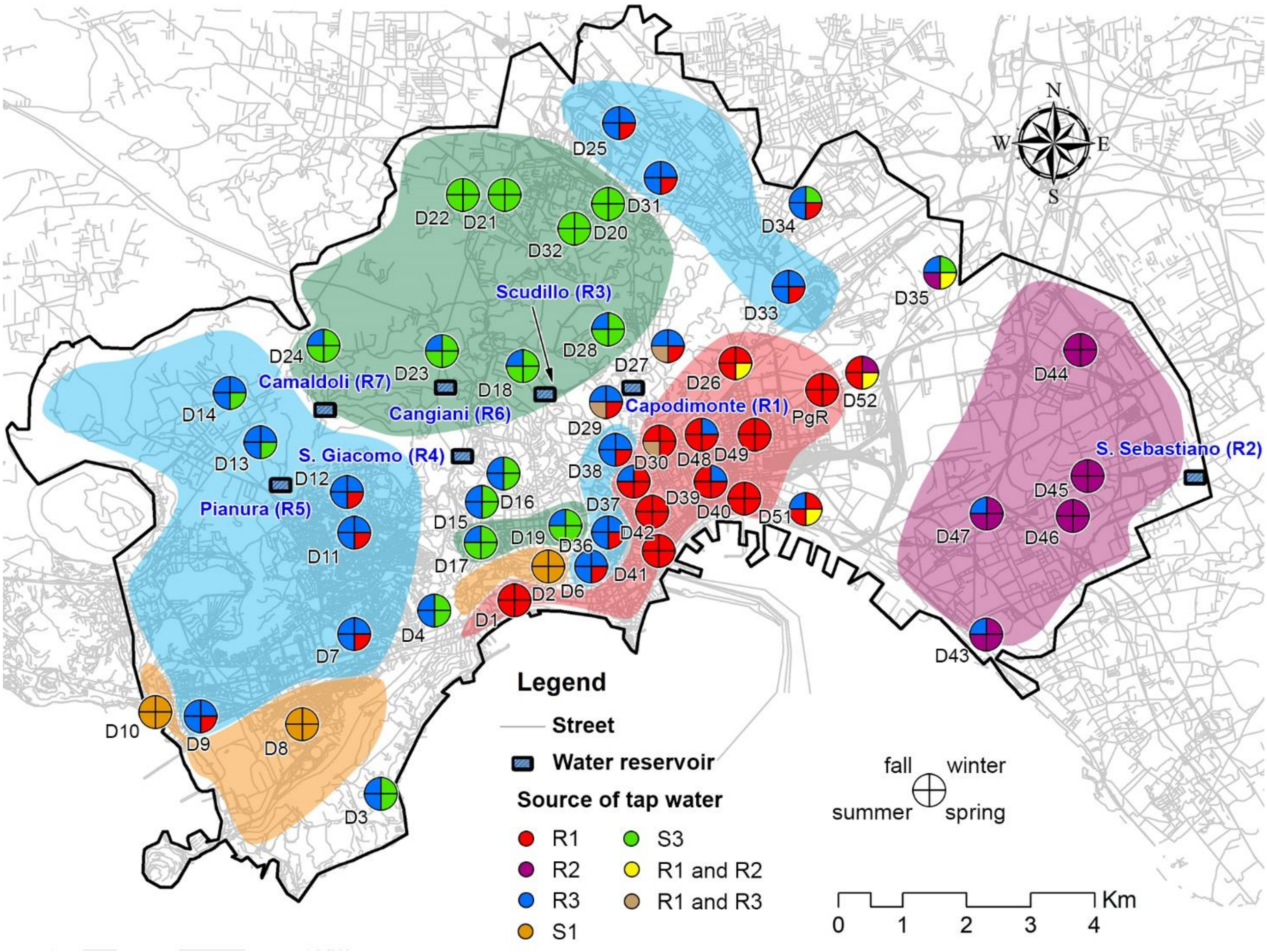


Fig. 16 Results of seasonal HCA indicating the city areas with the same source of tap water. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; and S3: Serino Aqueduct. Color of each city area should be matched with the same color in the legend to define the source of tap water.

According to the results and discussion in this chapter, the following aspects are considered to evaluate performance of HCA combined with clr-transformation: (1) D6, D51 and D52 are excluded from R1; (2) D17 and D29 are excluded from R3, but D6 and D33 are included in R3; (3) D33 and D35 are excluded from S2; and (4) D27 and D34 are excluded from S3. Average Ca^{2+}, Mg^{2+}, Na^{+}, Cl^{-}, SO_4^{2-}, NO_3^{-} and F^{-} contents are then compared in the samples before and after (represented without and with asterisk, respectively, in Fig. 17) implementing the abovementioned modifications. Figure 17a and b shows that the change in average chemical compositions of R1 and R3 is not clear. Hence, chemistry of the input waters is not represented for simplification. However, concentration of ions significantly increased in S2 after excluding D33 and D35. It is obvious that S2* is comparable with I3 and I4 indicating the highest contribution of these input waters and negligible importance of I5 in the period of this investigation (Fig. 17c). S3 is also plotted on I5 when D27 and D34 are ignored (Fig. 17d). Consequently, combination of HCA and clr-transformation works effectively with different sample sizes; however, capability of the approach can visually be verified when there are few samples.

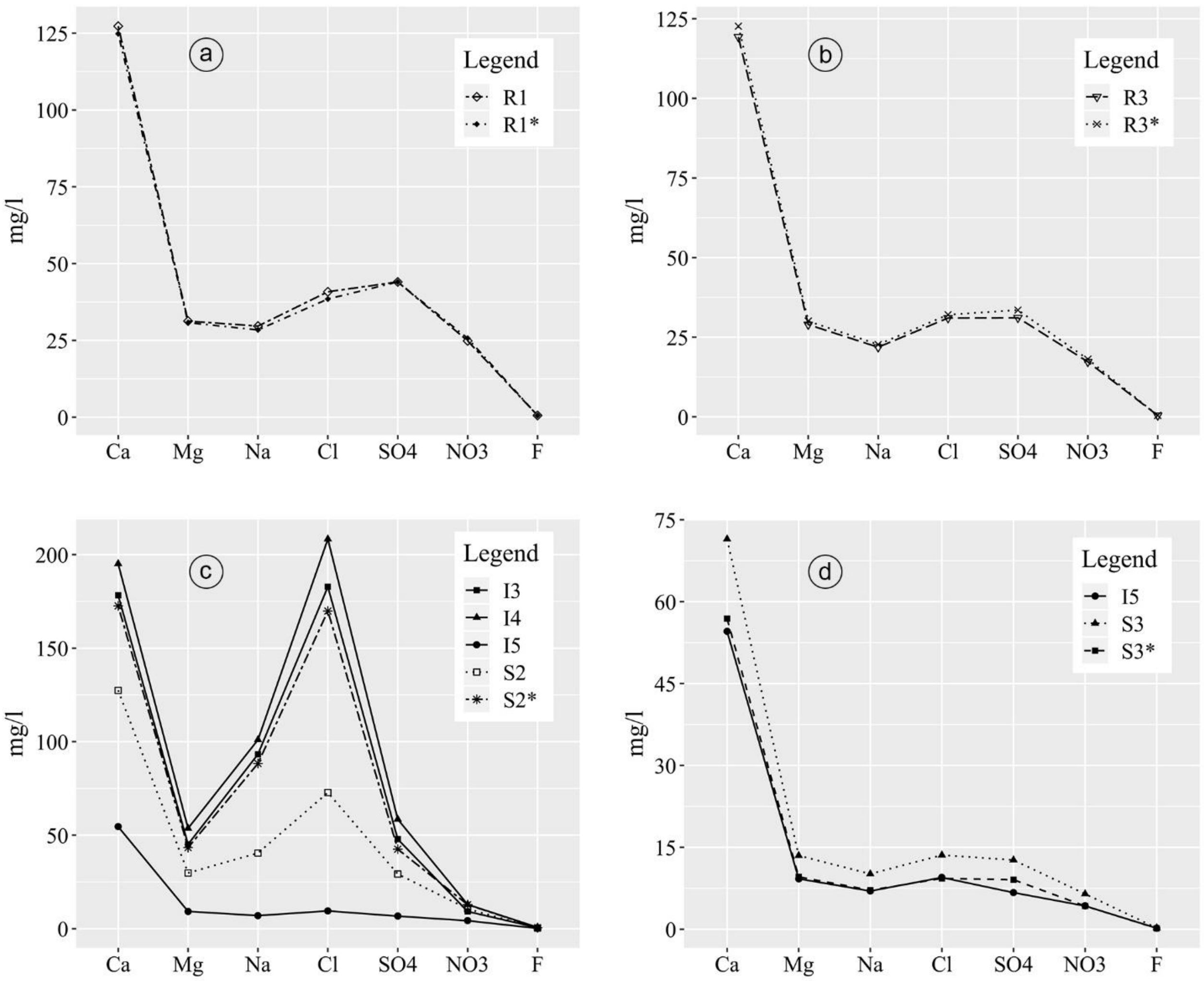


Fig. 17 The tap water composition before and after (without and with asterisk, respectively) implementing the HCA result: (a) R1: Capodimonte; (b) R3: Scudillo; (c) S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells, I3: Cancello well field, I4: Campano Aqueduct and Ponte Tavano wells, and I5: Serino Aqueduct; and (d) S3: Serino Aqueduct.

5. Conclusions

The obtained results show how the chemical data collected for periodical water quality checks can be used to assess the flow dynamics in a complex water distribution system fed by multiple supply sources with different geochemical fingerprints. Integrating the principle of CoDA and multivariate statistics for data treatment can be a valuable tool to seasonally monitor the prevalence of the different supply sources in a city like Naples where the urban distribution network has a historical stratified development and a large variety of materials used for constructing pipelines and urban reservoirs. Using the data acquired in different seasons to run the analysis in a wide timespan can represent an opportunity to have a comprehensive view of the distribution network dynamics which also favors the discovery of network interconnections not easily verifiable in the field. In short periods, development of multivariate analysis (especially HCA) using tap water chemistry can lead to a clearer time-specific view of the water distribution in the network together with more effective real-time interventions to locally improve water quality or to locate and manage pipeline failures and systemic water loss. The approach presented in this paper can be used in other complex urban water networks because of the limited cost and positive outcomes. Applying the proposed technique in more case studies and comparing the results with the hydraulic modelling techniques can encourage the scientific community and managers to consider it in future investigations. The performance of the technique can obviously be improved by increasing the sampling density.

Contribution of the Ph.D. candidate

Data visualization; Data interpretation; Writing the original draft; Revising the draft

CHAPTER 4

Marine Sediment Pollution in Pozzuoli Bay

This chapter was published in “Chemosphere” as a Research Article:

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. https://doi.org/10.1016/j.chemosphere.2021.129955.

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

A summary of the results was also presented in Goldschmidt2021 (Ebrahimi et al., 2021).

Marine Sediment Pollution in Pozzuoli Bay

Abstract

In the last decades, investigating geochemistry of sea sediments has been challenging in the eastern sector of Pozzuoli Bay, source of the metal(loid)s has been a matter of debate and the proposed origin of potentially toxic elements (PTEs) has been occasionally inconsistent. In this study, compositional data analysis (CoDA) was used because the results are independent of the measurement unit, the selected subgroup of elements and the order of chemicals in the dataset. The robust variant of principal component analysis (PCA) indicated that Hg, Cd, Cu, Pb and Zn were positively correlated with mud and organic matter in the sediments deposited in front of the former industrial site. Concentrations of these metals decrease along the cores and in the distal zone. Nevertheless, Al, As, V, Fe, Cr, Ni and sand form an association along the coast which strengthens with increasing distance from fumaroles in the proximal zone. It suggests that arsenic was mainly originated from the pyroclastic deposits of Phlegrean Fields and some of the seepages with a hydrothermal component, supported by low contribution of the variables in robust PCA of the sediments from distal zone. Therefore, this pioneering study suggests CoDA as a powerful tool for answering the long-lasting questions over sediment geochemistry in polluted areas.

1. Introduction

There has been a growing evidence of global environmental damage since the beginning of industrialization and urbanization (Beck et al., 2020). Volcanism and the related activities are also capable of bringing potentially toxic elements (PTEs) to the surface from deep within the Earth (Selinus, 2013). Coexistence of these features in a coastal area can lead to some complexities in geochemical investigations and characterizing source of pollution because sediments can be a source or a sink of chemicals (Keshavarzi et al., 2015; Beck et al., 2020). The restricted water circulation and presence of several industries along the coastline lead to high vulnerability of the Mediterranean Sea to anthropogenic activity. In the late 1980s, the pioneering studies reported industrial pollution of marine sediments in the eastern sector of Pozzuoli Bay (Damiani et al., 1987; Sharp and Nardi, 1987). It raised concerns towards ecological and human health risks of the pollutants and many researchers tried to find source and fate of chemicals by collecting beach sediments, seafloor surface sediments and sediment cores (ICRAM, 2005; Bergamin et al., 2005; Romano et al., 2004, 2009; Albanese et al., 2010; Trifuoggi et al., 2017; Armiento et al., 2020; Daliri et al., 2020; Giglioli et al., 2020). Bergamin et al. (2005) documented positive correlation of Cu, Cr, Mn and Zn with the number of deformed benthic foraminifera. Significant contamination of Cd, Cu, Hg, Mn, Pb and Zn in the seafloor surface sediments collected about 15 years after closure of the local industrial plants indicated that the industrial site is still a source of pollution (Romano et al., 2009).

In the last decade, Albanese et al. (2010) illustrated that the metal-rich groundwater which mainly originates from hydrothermal activity of Phlegrean Fields has carried the pollutants from the coastal landfill to the marine environment. The highest As concentrations showed a strong spatial correlation with the submarine thermal springs. Arsenic contents decreased with increasing seawater depth (Somma et al., 2016) because hot hydrothermal fluids probably limit As precipitation (Albanese et al., 2010). Trifuoggi et al. (2017) dredged seafloor surface sediments of Pozzuoli Bay along five transects perpendicular to the coastline and showed that the average concentrations of As, Cd, Hg, Pb and Zn were two to twenty times greater than the Italian regulatory guidelines. They concluded that the insignificant correlation between the mean grain size and PTEs (i.e., As, Cd, Cr, Cu, Hg, Ni, Pb and Zn) excludes terrigenous sediments and weathering products as their main sources.

Recently, Armiento et al. (2020) showed that As, V, Cr and Ni were controlled by geogenic (volcanism and the related fumaroles for As and V, but minerals for Ni and Cr) and anthropogenic factors. Nevertheless, Giglioli et al. (2020) demonstrated that pyroclastic rocks, groundwater, seepages, fumaroles and industrial activity have introduced As to the coastal environment. While wave hydrodynamics affects spatial distributions of the metals originated from the coastal landfill (i.e. Cd, Cu, Hg, Pb, Zn and Fe) together with As, it does not influence geochemical distributions of Cr, Ni and V mostly concentrated in the sediments deposited offshore and NW Nisida island. Because sediment toxicity is not mostly representative of sediment chemistry, Morroni et al. (2020) used an integrated approach and identified a “Slight” to “Moderate” level of risk, with the highest near the plant.

Regarding the abovementioned studies, the underwater fumaroles, submarine groundwater discharge, minerals, extensive industrial activity in the past, ineffective disposal of the industrial wastes in the coastal landfill and the seepages were reported as the factors governing sediment geochemistry in the study area. It indicates that there has been a long-lasting debate between researchers about source of some PTEs (particularly As) in the marine sediments due to the complex geochemical environment. Compositional data analysis (CoDA) is getting popular because of its potential to avoid doubtful results. Separation of the geochemical signals of grain size, anthropogenic contamination and possible post-depositional alteration in fluvial sediments of NW Czech was facilitated by utilizing CoDA tools (Álvarez-Vázquez et al., 2020). In another study, Dominech et al. (2020a) investigated cadmium in the sediments collected along Sabato river (south Italy) and reveled capability of CoDA in showing the relative geochemical information at catchment basin scale. The extent of mineralization, weathering, diagenesis, contamination and a combination of these factors can be effectively detected following log-ratio transformation of the analytical data obtained from stream sediments (Wang et al., 2014; Liu et al., 2016, 2018). The idea behind CoDA was also applied for paleoenvironmental reconstruction of sedimentary environments (Flood et al., 2016; Razum et al., 2020).

To the best of our knowledge, log-ratio transformations were only utilized by Wang et al. (2015) to generate single-element geochemical maps in marine sediments. Therefore, based on the CoDA framework, this is the first work that presents the results of multivariate statistical analysis of PTEs in seafloor surface sediments and core sediments. It is expected that considering compositional nature of data leads to some useful information about the origin of elements and reveals the capabilities of this approach for investigating geochemistry of polluted sea sediments. Furthermore, a data transformation will be proposed to enable fractal analysis for geospatial analysis of the negative values obtained after log-ratio transformation.

2. Study area

The Bagnoli urban district (Fig. 1) in the city of Naples is located on the eastern coast of Pozzuoli Bay (Tyrrhenian Sea, Italy). For about 100 years, production of steel, asbestos, cement and fertilizer in an industrial site established near the urban area strongly polluted the environment. It is worth mentioning that the steel plant was one of the most important Italian steelworks in the 20th century which reached a maximum of 2×10^6 million tons of steel production in 1960. The modifications of human to the natural environment have been one of the key factors in environmental crisis in the study area. In 1930, two piers were constructed for ship berthing (Fig. 2). The raw materials such as coal and iron ore were unloaded in the northern one although the products were loaded in the southern one. Marine hydrodynamics was accordingly changed, and sediment transport capacity has significantly decreased in front of the northern pier. Five years later, Nisida island was connected to the mainland by a stone bridge which also altered the water circulation pattern and sea sediment movement. In order to construct buildings, store coal and develop the industrial activity, the area between the piers was partly filled with contaminated soil of the industrial site in the early 1960s (Fig. 1) (Romano et al., 2009; Arienzo et al., 2019; Armiento et al., 2020). The shoreline dynamics generally depends on the influence of wind and tide on the emerged beach together with sediment mobilization by wave motion and marine currents in the submerged beach. The winds and waves that approach from southwest and west lead to a NW-SE drift by longshore currents and control the littoral morphodynamics. The effective fetch and depth of closure calculated for 1951-2009 period are 280 km and 7.5 m, respectively (Arienzo et al., 2019).

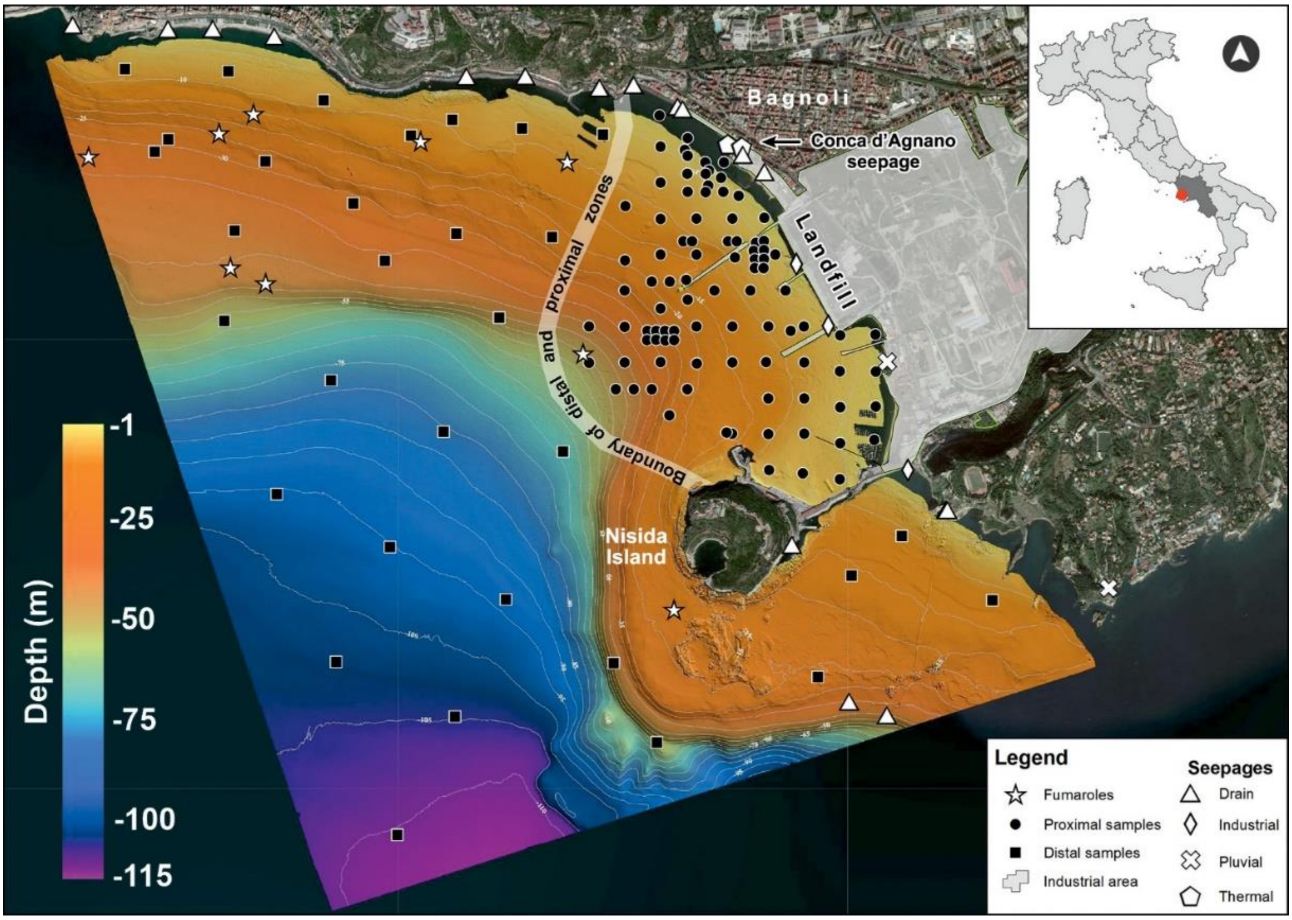


Fig. 1 Location of the study area, bathymetry of the coastal environment, the sediment samples of proximal and distal zones (PZN and DZN, respectively) and location of seepages and fumaroles.

On the other hand, the study area is a part of the Campania Volcanic Zone (Rolandi et al., 2020) and lies within the active volcanic area of Phlegrean Fields caldera (Fig. 2) formed 15 ka B.P. by a large ignimbrite eruption named Neapolitan Yellow Tuff (De Natale et al., 2016; Troise et al., 2019). Neapolitan Yellow Tuff is covered with a sequence of interdigitated pyroclastic deposits produced by recent volcanic and tectonic activities and then lacustrine and palustrine sediments. The last eruption occurred in 1538 CE, but the intense hydrothermal circulations have resulted in surface manifestations of hot springs, steam-heated pools, together with submarine (Fig. 1) and subaerial fumaroles. The ascending hydrothermal fluids bring considerable loads of PTEs such as arsenic to the surface (Aiuppa et al., 2006).

The abovementioned discussion reveals that concentrations of PTEs are regulated with anthropogenic and geogenic features. Although closure of the industrial site occurred in 1990s, the area which poses a high environmental risk was included in the list of polluted sites of national interest (SIN) for monitoring and recovering the area. Remediation programs were started in 1996 and extended to the marine sediments in front of the brownfield five years later (Albanese et al., 2010). ABBaCo was the latest project which aimed to (ABBaCo Project, 2018): (1) upgrade the information about environmental quality of Bagnoli industrial site; (2) characterize the pollution sources; and (3) propose appropriate remediation strategies.

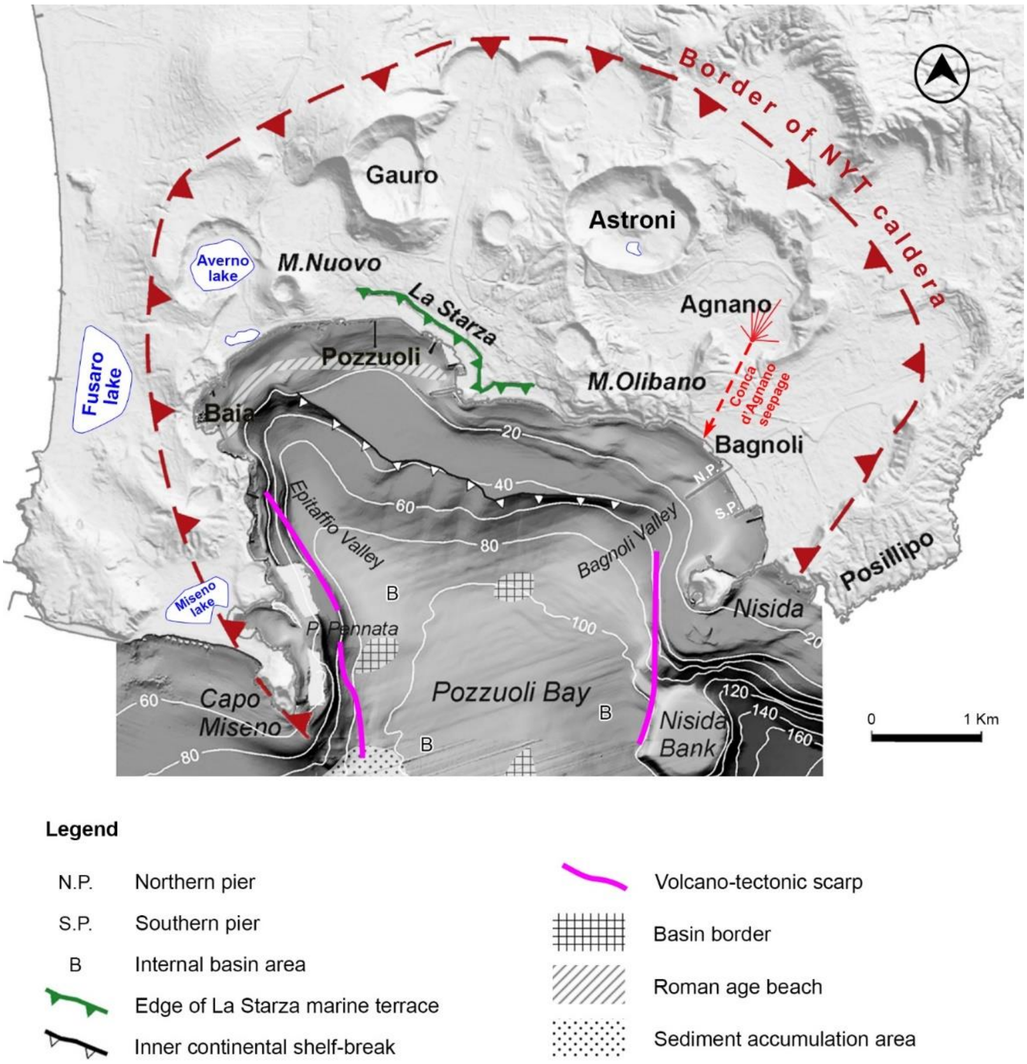


Fig. 2 A small-scale map of the study area showing the Neapolitan Yellow Tuff (NYT) caldera and different marine environments in Pozzuoli Bay (after De Pippo et al. (1988) and Sacchi et al. (2014)).

3. Materials and methods

3.1. Sampling and chemical analysis

Sampling campaigns were planned in May and November 2017 to collect sediment samples from the proximal (near the Bagnoli industrial area) and distal (far from the industrial area) zones (hereafter PZN and DZN, respectively; Fig. 1) regarding the preliminary investigations of the Italian Higher Institute for Environmental Protection and Research (ISPRA). In the PZN, 93 sediment cores were drilled using a vibro-corer with 10 cm inner diameter. However, a Van Veen grab was used to collect 32 samples from the upper 20 cm in the DZN (Fig. 1). About 1 kg of sediment was placed into a plastic bag and stored at approximately 4 °C until transfer to the laboratory. Coordinates of the sampling stations were obtained via a differential global positioning system.

Laboratory sampling and sediment pre-treatment were performed following the Ministerial Decree 7/11/2008. Each sediment core was accordingly sub-sectioned into 0-50, 50-100, 100-150, 150-200 and 200-

300 cm intervals depending on the core depth. The aliquots for metals/metalloids (except for Hg) were oven-dried at 105 °C for 24 h whilst the ones for Hg were air-dried. All samples were subsequently homogenized and stored at 4 °C before laboratory analyses according to the guidelines and directives for assessment of marine pollution in SINs (Legislative Decree, n. 152/06). For grain size analysis, the sediments were successively treated with H_2O_2 solution, washed with distilled water and dried at 40 °C prior to weighing. Samples were then wet-sieved using a 63-µm sieve, and the coarser fraction was further separated into different grain sizes by passing them through ASTM series sieves with meshes ranging between –2 and +4 phi (Romano and Gabellini, 2001). Finally, the muddy (<63 µm), sandy (63 µm-2 mm) and gravelly (>2 mm) fractions were calculated for the purpose of this study.

Total concentrations of elements were measured in the <2 mm grain size by microwave-assisted acid digestion procedure following the EPA 3052 method. The analyses of the metal(loid)s were performed through the following protocols and instruments: (1) EPA 6020b for As, Cd, Cr, Cu, Ni, Pb, V and Zn determination with inductively coupled plasma-mass spectrometry (PerkinElmer Elan-6100); (2) EPA-6010d for Al and Fe measurement with inductively coupled plasma-optical emission spectrometry (PerkinElmer Optima-2000DV); and (3) EPA 7473 for Hg detection with the Automatic solid/liquid Mercury Analyser (FKV AMA-254). The computed quantitation limits for dry weight were as follows (mg/kg): 0.05 for As, V and Pb, 0.03 for Cd, 0.5 for Cr, 0.3 for Cu and Ni, 1.0 for Zn, 50 for Al, 5 for Fe and 0.005 for Hg. In order to check accuracy of the analytical determinations, BCSS-1, MESS-4 and PACS-3 marine sediment standard reference materials were analyzed using the procedures applied to the sediment samples and the obtained recovery rates ranged from 80 to 118% (Table 1).

Table 1 Recovery (%) of each element regarding the analyzed marine sediment standard reference materials.

Element		**Al**	**As**	**Cd**	**Cr**	**Cu**	**Fe**	**Hg**	**Ni**	**Pb**	**V**	**Zn**
Recovery (%)	**BCSS-1**	86	95	108	85	103	86	90	98	95	98	107
	MESS-4	84	94	110	95	101	80	95	97	100	94	118
	PACS-3	88	95	98	90	89	85	-	96	90	98	100

3.2. Compositional data analysis (CoDA)

In geochemical and environmental datasets, variation of most parameters should be assessed relative to the other variables because only ratios between parts are informative. This is well explained by the fundamentals of CoDA in which a log-ratio methodology transforms the compositional data from their original sample space to an unrestricted real space (Filzmoser et al., 2018). The proposed family of log-ratio transformations includes the additive log-ratio (alr), centered log-ratio (clr) and isometric log-ratio (ilr) transformations (Aitchison, 1982; Egozcue et al., 2003). The values obtained from the first two log-ratios are subject to some restrictions in standard statistical techniques whilst the ilr-transformed data can be investigated by standard univariate and multivariate statistical approaches (Lloyd et al., 2012). In the present study, the elements and the grain size fractions were considered as two different sets of compositional data.

3.2.1. Spatial distribution of single variables: raw data vs. clr-transformed data

In order to obtain a general insight into the coastal sediment geochemistry, the interpolated maps of the raw data and the clr coordinates were generated (Fig. 3). This transformation has been used to decipher the geochemical patterns in sediments by several researchers (e.g., Liu et al., 2016; Dominech et al., 2020a, 2020b). The clr transformation for a composition $\mathrm{x} = [\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{D}] \in \mathrm{S}^\mathrm{D}$ is the transformed data $\mathrm{y} \in \mathrm{R}^\mathrm{D}$ with (Lloyd et al., 2012):

$$\mathrm{y} = \mathrm{clr}(\mathrm{x}) = \left(\ln \frac{\mathrm{x}_1}{\sqrt[\mathrm{D}]{\prod_{\mathrm{i}=1}^{\mathrm{D}} \mathrm{x}_\mathrm{i}}}, \ln \frac{\mathrm{x}_2}{\sqrt[\mathrm{D}]{\prod_{\mathrm{i}=1}^{\mathrm{D}} \mathrm{x}_\mathrm{i}}}, \ldots, \ln \frac{\mathrm{x}_\mathrm{D}}{\sqrt[\mathrm{D}]{\prod_{\mathrm{i}=1}^{\mathrm{D}} \mathrm{x}_\mathrm{i}}} \right) \tag{1}$$

Computing the clr coefficients is not complicated and assigning each original compositional part to a clr coordinate relatively facilitates interpretation of the transformed values. However, this transformation is sub-compositionally incoherent which may result in misinterpretations when including a subset of variables in the denominator (Razum et al., 2020). Another restriction of the clr coefficients is that they sum to zero and their covariance and correlation matrices are singular (Buccianti et al., 2006). Thus, the ilr transformation (explained more in the following sections of this chapter) was also considered for validation of the geochemical maps produced following clr-transformation and for identification of the main processes regulating geochemistry of PTEs.

The geochemical mapping was performed by the inverse distance weighted (IDW) method (Fig. 3). The search radius was set to 300 m for proximal zone and 1000 m for distal zone, and the 4 nearest sample points were used to generate the maps with output cell size of 10 m. Because distribution of geochemical data has a fractal nature, the interpolated surfaces were then reclassified through the concentration-area (C-A) model. In the log-log plot for the concentration-area relation (the C-A plot) of each element, the vertical axis shows cumulative area of the pixels with concentrations above ρ (i.e. A(ρ)), and the horizontal axis represents the measured concentrations (ρ). The sudden changes in slope of the fitted straight-line segments are then considered and the corresponding values of ρ are used as cut-off scores to classify pixel values into different groups (Cheng et al., 1994). This approach can preserve enrichments regarding spatial associations and local singularities in the geochemical data at local to continental scales (Lima et al., 2003, 2008; Albanese et al., 2007, 2015; Cicchella et al., 2015; Thiombane et al., 2018b; Jordan et al., 2018).

Because a non-negative dataset is required for mapping, a data treatment was proposed in this chapter to deal with the negative clr coefficients (Fig. 3). First, the variables (i.e. Al, Fe and Zn) with positive clr values ($clr.element$) were scaled ($clr_{scl1}.element$ and $clr_{scl2}.element$) using equations 2 and 3 in Table 2. Then, the spatial distributions of $clr.element$, $clr_{scl1}.element$ and $clr_{scl2}.element$ were prepared and reclassified according to the fractal approach. The upper and lower values of each class were subsequently back-transformed to the original $clr.element$ by the equations 4 and 5 (Table 2). It was illustrated that applying the proposed formulas to scale the clr coordinates does not significantly affect the spatial patterns (Figs. 4-6). Therefore, whenever a negative $clr.element$ is present in a variable, Eq. (2) is utilized to obtain the threshold concentrations by fractal analysis which will be back-transformed to the original clr coefficients via equation 4. The abovementioned approach was also utilized for mapping grain size fractions, principal component scores and standardized residuals which will be further discussed in the following sections of this chapter. The MIDW mapping was performed using the ArcFractal Plugin in ArcGIS 10.4.1 software.

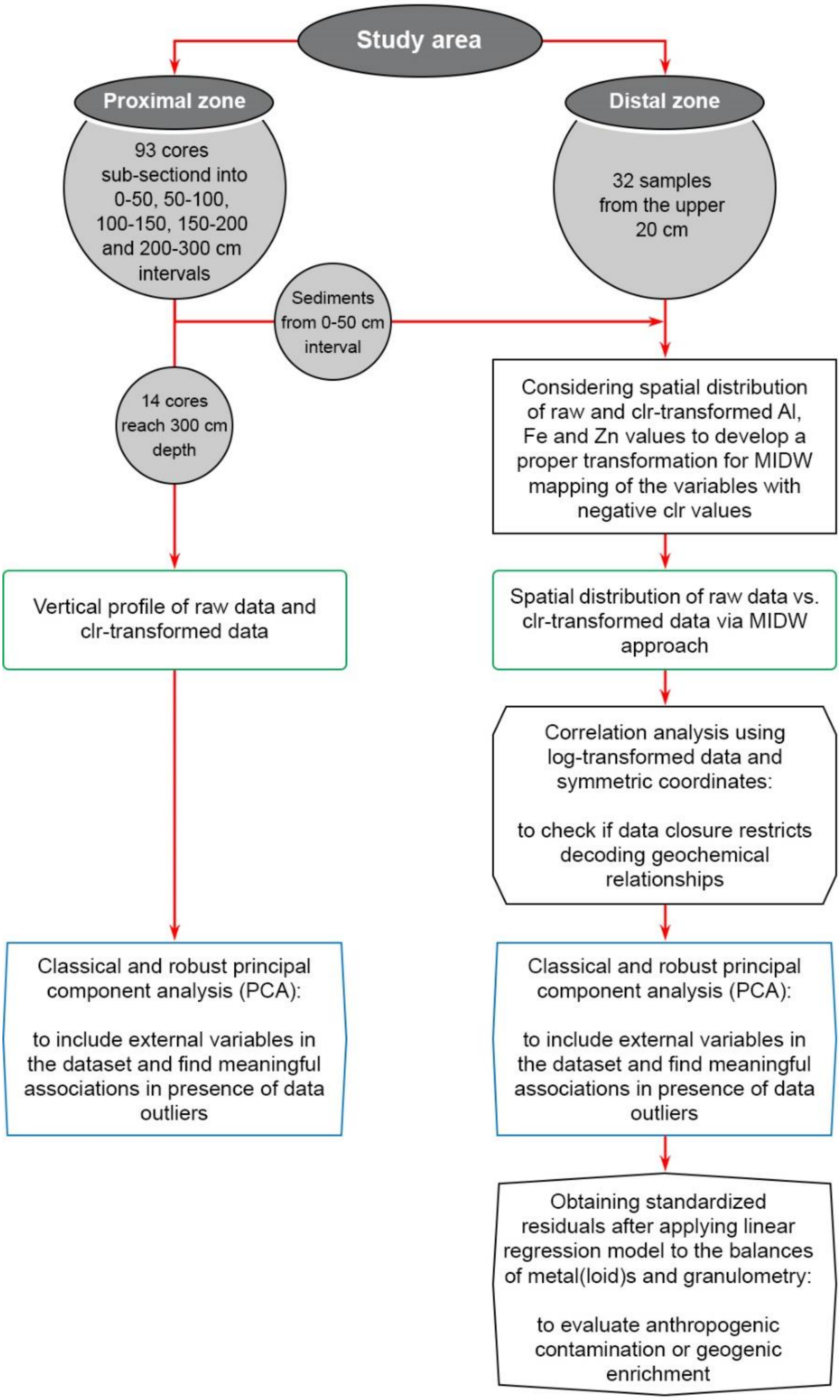


Fig. 3 The flow chart of data treatment in the present chapter.

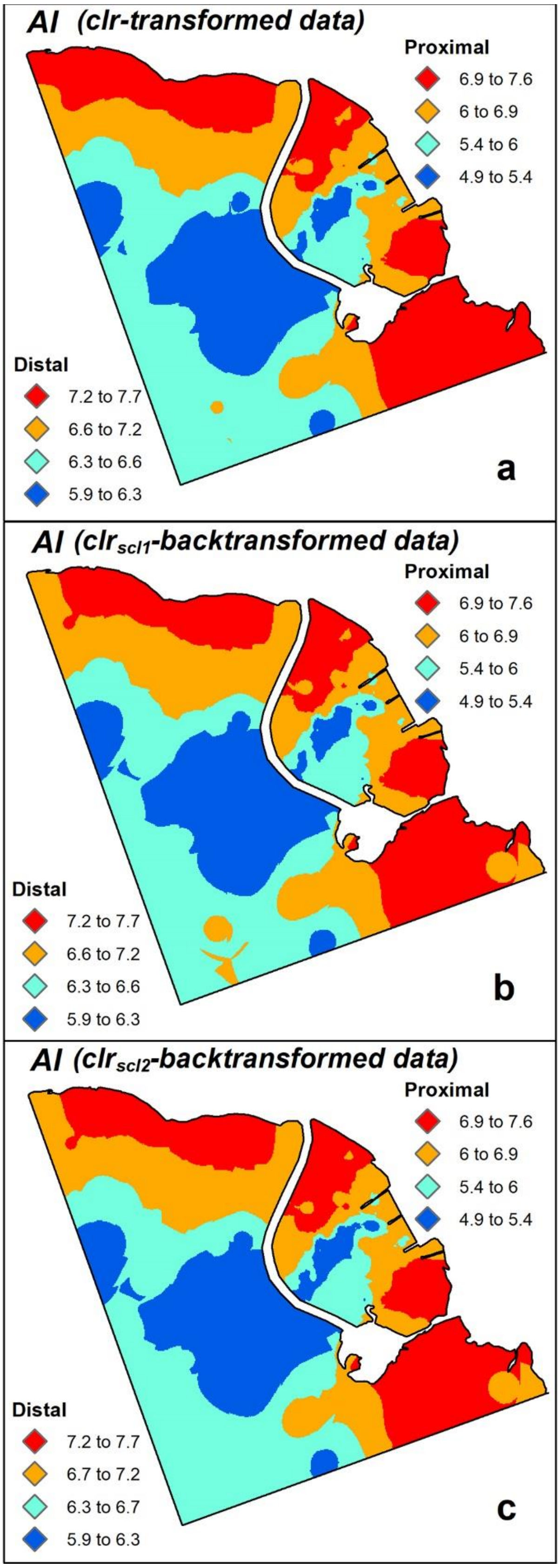


Fig. 4 Spatial distribution of (a) clr-transformed, (b) clr_{scl1}-backtransformed and (c) clr_{scl2}-backtransformed values of aluminum. Regardless of the applied data transformations, the patterns remained almost stable.

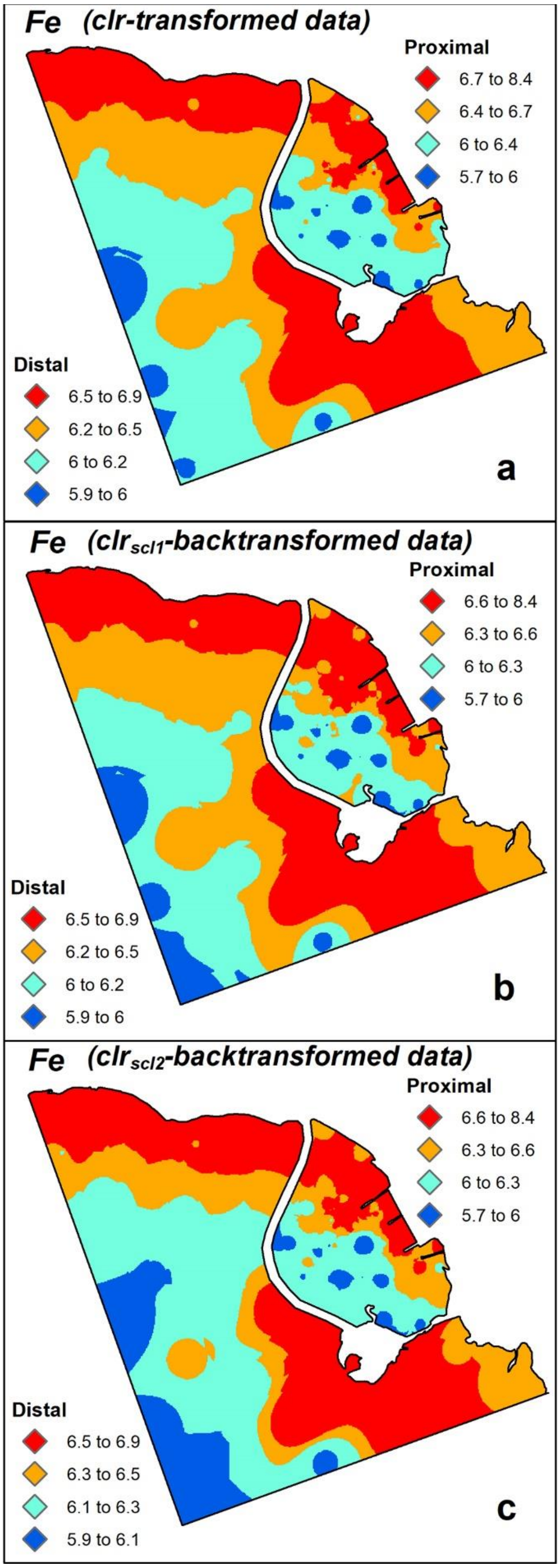


Fig. 5 Spatial distribution of (a) clr-transformed, (b) clr$_{scl1}$-backtransformed and (c) clr$_{scl2}$-backtransformed values of iron. Regardless of the applied data transformations, the patterns remained almost stable.

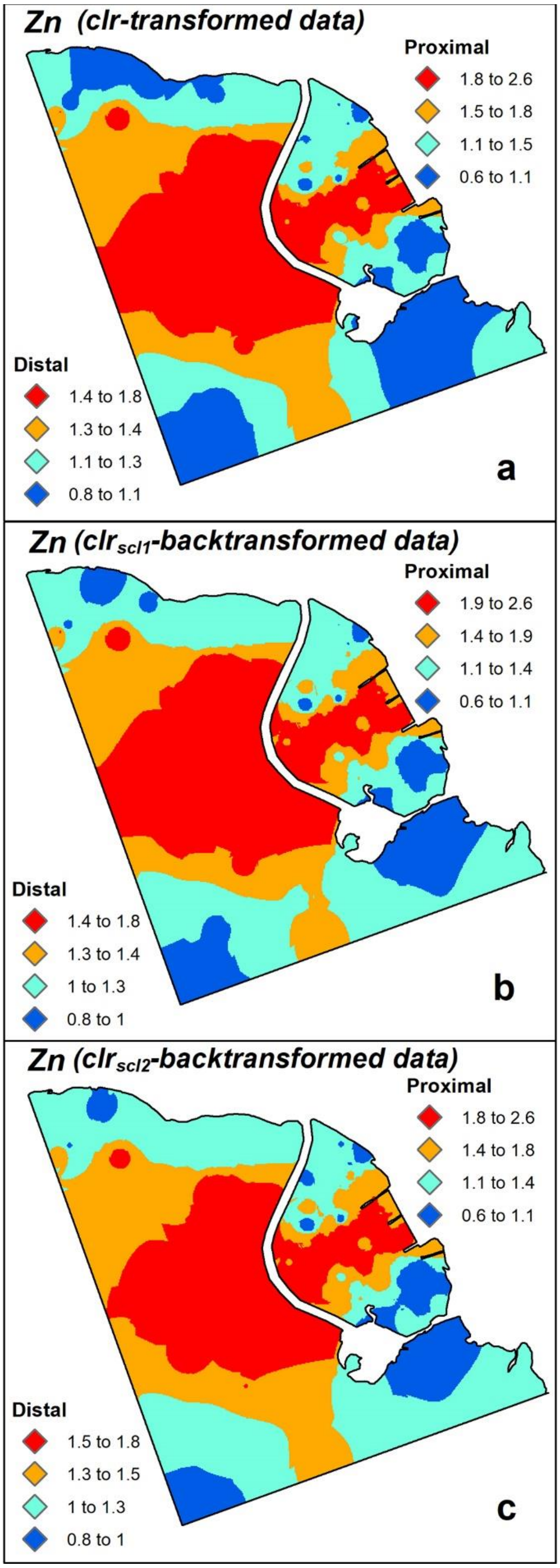


Fig. 6 Spatial distribution of (a) clr-transformed, (b) clr$_{\text{scl1}}$-backtransformed and (c) clr$_{\text{scl2}}$-backtransformed values of zinc. Regardless of the applied data transformations, the patterns remained almost stable.

3.2.2. Correlation analysis: the log-transformed data vs. the symmetric coordinates

At the end of nineteenth century, Pearson (1897) explained the problems with correlation analysis applied to the parts of a composition and demonstrated that obtaining reliable results depends on capturing the relative information of compositional parts. The clr transformation is not suitable for this purpose due to the drawbacks discussed before. Pivot coordinates (a special choice of the ilr transformation) are an alternative, but the direct relationship between the original variables and the transformed values will be lost (Reimann et al., 2017). To overcome the shortcomings of clr and ilr transformations and evaluate the relationship between two compositional parts, construction of the symmetric coordinates as a special case of the ilr transformation is advantageous. It preserves the dominance of both parts in a composition by generating two pivot coordinate systems in a symmetric manner for a D-part composition represented as $x = (x_1, \ldots, x_D)'$ using the following formula (Kynčlová et al., 2017):

$$z_i = \sqrt{\frac{D-i}{D-i+1}} \ln \frac{x_i}{\sqrt[D-i]{\prod_{j=i+1}^{D} x_j}} \qquad \text{for} \qquad i = 1, \ldots, D-1 \qquad (6)$$

The balance z_1 evidently contains all the relative information of the part x_1 with respect to the remaining compositional parts because this part is excluded from the other coordinates. The same interpretation cannot be extended to z_2 and x_2 due to exclusion of the first part in this balance. Nonetheless, a symmetric form of z_1 and z_2 would be an acceptable candidate for the correlation between relative contributions of x_1 and x_2 in a given composition due to the exclusive position of the parts of interest (x_1 and x_2) in the respective coordinates (Kynčlová et al., 2017). A similar construction for different compositional parts can be obtained by permuting the parts in Eq. 6.

Because the classical log-transformation of single variables prior to evaluating the relationship between parameters is still common, the log-transformed data and the symmetric coordinates were utilized to conduct correlation analysis and demonstrate the importance of dealing with data closure (Fig. 3). The R package “robCompositions” (Filzmoser et al., 2018) was used for transforming the data and generating the heatmaps of Spearman correlation.

Table 2 Scaling the computed clr values ($clr.element$) to obtain positive values for MIDW interpolation. The cut-off scores were subsequently back-transformed to the original $clr.element$.

	Scaling the *clr.element*		**Back-transforming to the *clr.element***	
All variables	$\mathrm{clr_{scl1}.element = clr.element - Min_{clr.element} + 0.1}$	Eq. (2)	$clr.element = \mathrm{clr_{scl1}.element + Min_{clr.element} - 0.1}$	Eq. (4)
Variables with *clr.element* > 0	$\mathrm{clr_{scl2}.element = clr.element + Min_{clr.element} + 0.1}$	Eq. (3)	$clr.element = \mathrm{clr_{scl2}.element - Min_{clr.element} - 0.1}$	Eq. (5)

3.2.3. Principal component analysis (PCA): the classical vs. robust approaches

Aitchison and Greenacre (2002) suggested the clr transformation for dealing with data closure in the context of PCA. The covariance structure of compositional data is significantly biased without an appropriate data transformation and in presence of data outliers (Maronna et al., 2006) which leads to doubtful results of multivariate statistical analysis. Robust estimation of the covariance matrix was proposed as a solution in the robust PCA (Filzmoser, 1999). Contrary to classical PCA (Aitchison and Greenacre, 2002), it does not work with clr-transformed data because a full rank data matrix is usually needed for robust covariance estimators. Therefore, Filzmoser et al. (2009) proposed applying the ilr transformation prior to robust PCA and subsequently back-transforming the resulting scores and loadings to the clr space for better interpretation (Fig. 3). The results can be visualized via biplots which are useful exploratory tools for two-dimensional graphical representation of observations and variables (Gabriel, 1971). The form biplot helps to assess the display of individuals (i.e. scores of principal components), but the covariance biplot favors the display of variables. In the latter, length of the rays (arrows) indicates standard deviation of the respective coordinates whilst the cosine of the angle between two rays approximates the correlation coefficient between the corresponding variables. The distance of a data point from the origin of a biplot reveals the Mahalanobis distance of the point from the mean, being useful for detecting outliers as well (Filzmoser et al., 2018). In a biplot, when vertices of three rays occur approximately on a straight line, it probably reveals one-dimensional variability amongst the components (Aitchison et al., 2002).

3.2.4. Sediment pollution

Although CoDA is a powerful tool for geochemical investigation (Thiombane et al., 2018b; Dominech et al., 2020a, 2020b), the currently available pollution indices do not address the compositional nature of data. To bridge this gap, balances which are a family of the ilr coordinates were constructed through sequential binary partitioning. It helps to easily interpret the relative information about two groups of parts that cause the major effects within a composition. Specifically, the PCA results in this chapter were considered for selection of two non-overlapping groups of compositional parts. The first group was coded with "+" and consisted of p parts whilst the second one was coded by "–" and composed of m parts. Thus, $i_1, i_2, \dots, i_p$ and $j_1, j_2, \dots, j_m$ represent the first and second groups of compositional parts, respectively. Then, the corresponding balance ($\tilde{z}$) was computed according to the following formula (Egozcue and Pawlowsky-Glahn, 2005):

$$\tilde{z} = \sqrt{\frac{pm}{p+m}} \ln \frac{(x_{i_1} x_{i_2} \dots x_{i_p})^{\frac{1}{p}}}{(x_{j_1} x_{j_2} \dots x_{j_m})^{\frac{1}{m}}} \qquad (7)$$

Two balances were separately generated by metal(loid)s and granulometry as different sets of compositional data. Linear regression model was subsequently applied to the balances for modelling the relationship between metal(loid)s and granulometry as dependent and independent variables, respectively. The residuals were divided by their standard error to obtain the standardized residuals which can be indicators of anthropogenic contamination or geogenic enrichment (Fig. 3) (Petrik et al., 2018b). These values were finally mapped via MIDW approach as described in the section 3.2.1. The linear regression model was performed through the R package "robustbase" (Maechler et al., 2020).

4. Results and discussion

4.1. Geochemistry of seafloor surface sediments

This section refers to the sediments from 0 to 50 cm in the PZN and 0 to 20 cm in the DZN.

4.1.1. Physicochemical characteristics of the sediments

The descriptive statistics of grain size fractions and organic matter (OM) contents of the samples are shown in Table 3. In the PZN, the proportion of gravel, sand and mud range from 0-22%, 26-99% and 1-70%, respectively. On the other hand, variation of sediment grain size is higher in the DZN: 0-55% for gravel, 9-99% for sand and 0-91% for mud. On average, sand is the dominant fraction. However, higher variation of mud (CVR > 70%) and gravel (CVR > 85%) in both datasets might explain the role that wave plays in transporting sediments (Table 3). The sediment type (Schlee, 1973) is mainly sand, silt and clay, but some samples are gravelly sediment (Fig. 7). It is worth mentioning that because elements were determined in the <2 mm size fraction, the gravel was only used for sediment type classification. According to Fig. 8, sand is mainly deposited along the coast whilst mud is more frequent (sometimes the major fraction) in the low-energy environment in front of Bagnoli landfill (De Pippo et al., 2002). The clr-transformation of the data does not change the spatial pattern of coarse and fine sediment fractions (Fig. 8a-d).

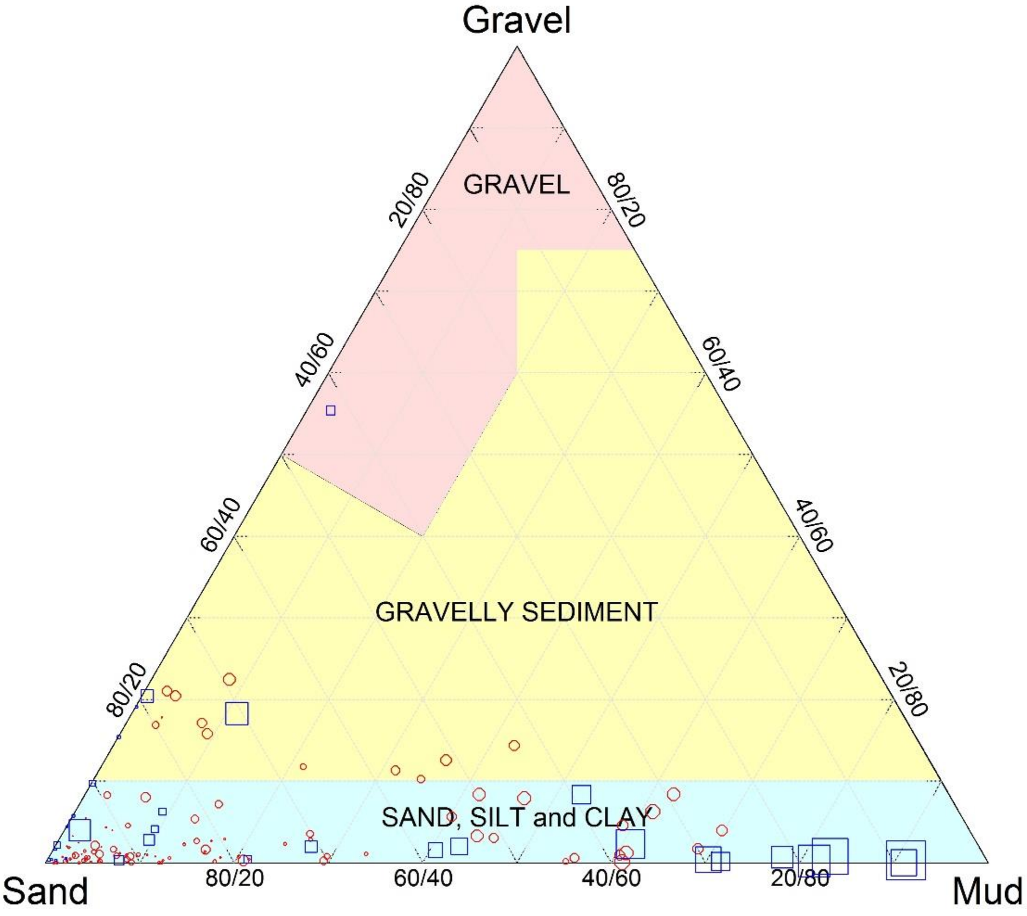


Fig. 7 The ternary diagram for classification of sediments based on gravel, sand and mud fractions (Schlee, 1973). The circles and squares represent the sediments from proximal (0-50 cm) and distal (0-20 cm) zones, respectively. Size of the symbols relatively shows the distance to the coastline.

The median OM content of the sediments from PZN (2.18%) is approximately twice that of DZN (0.92%). The highest OM values occur near the Bagnoli brownfield site and the levels decrease towards the open sea and the shoreline (Fig. 8e). The fine-grained sediments and OM can be important metal carriers (Ebrahimi, 2013; Sciarrillo et al., 2020), which will be further investigated in the following sections of this chapter.

4.1.2. Metal(loid) contents of the sediments

Statistical summary of the analyzed metal(loid)s is presented in Table 3. Considering the 95th percentile, concentrations of the elements are greater in the PZN except for Ni, Cr and V with almost equal levels in both zones. The average contents of the elements follow the decreasing order of Fe > Al > Zn > Pb > V > As > Cu > Cr > Ni > Cd > Hg in the PZN and Al > Fe > Zn > Pb > V > As > Cr > Cu > Ni > Cd > Hg in the DZN. Contrary to Al, Fe, As, Ni, Cr and V, the CVR values of Hg, Cd, Pb, Zn and Cu are above 60% in the PZN which likely suggest presence of unusual processes (Reimann et al., 2008). The CVR values of Cd, Pb and Zn decrease in the DZN while those of Hg and Cu change insignificantly.

Evidently, Fig. 9a to l indicates that there is a considerable difference between spatial distribution of raw data and the clr coordinates of As, V, Al, Fe, Cr and Ni. Thus, it might stress the importance of opening the geochemical data before geospatial analysis. Regarding the MIDW maps produced following transformation (Fig. 9a, c, e, g, i and k), the higher “clr.As”, “clr.V”, “clr.Al”, “clr.Fe”, “clr.Cr” and “clr.Ni” values mainly occur along the coastline and do not coincide with the industrial activities. However, spatial patterns of the other chemicals remained almost invariable after log-ratio transformation. In PZN, the sediments characterized by almost high contents of Hg, Cd, Pb, Zn and Cu are deposited in front of the local industrial site and the highest concentrations occur in greater water depth (Fig. 9m to v). This geochemical pattern can be explained by wave turbulence in front of the brownfield which encourages suspension and transportation of sediments from the emerged and submerged beaches (i.e. from 4.3 m a.s.l. to 6.5 m b.s.l.) due to run-up and depth of closure (Arienzo et al., 2019). Extension of the contaminated plume in the DZN can also be verified through the elevated clr values of these PTEs.

4.1.3. The statistical significance of data transformation

The Spearman correlation between the elements is presented in the proximal and distal sediments (Fig. 10). It is evident that, in agreement with the aforementioned discussion, CoDA clustered the variables into two distinct groups (Fig. 10a and c): (1) Al, As, Cr, Ni, V and Fe; and (2) Hg, Pb, Zn and Cu. Cadmium is correlated with the second group in the sediments near the industrial site, but it is associated with the first one in the sediments far from Bagnoli. In addition, Cr and Ni revealed weaker correlation coefficients with Al, As, V and Fe in the DZN. However, the majority of elements are positively correlated together in the right panels constructed via log-transformed data (Fig. 10b and d) which prevents effective data interpretation. The compositional approach was consequently used in this investigation for reliable data mining and informative pattern recognition.

Table 3 Descriptive statistics of metal(loid)s and the physicochemical parameters in the sediments collected from proximal (0-50 cm depth) and distal (0-20 cm depth) zones. Spatial distribution of the samples is illustrated in Fig. 1. The elements are in mg/kg and the rest are in percentage.

	Al	**As**	**Cd**	**Cr**	**Cu**	**Fe**	**Hg**	**Ni**	**Pb**	**V**	**Zn**	**Gravel**	**Sand**	**Mud**	**OM**
Proximal zone (N = 93): 0-50 cm															
Min	27827.0	28.2	0.28	13.0	8.6	21579.0	0.02	4.3	25.3	42.0	93.6	0.00	26.26	0.87	0.25
P5	39964.4	41.4	0.34	16.1	11.7	34467.6	0.04	6.7	60.3	83.4	121.8	0.03	37.28	1.37	0.30
P25	53300.0	56.3	0.45	20.8	16.8	47112.0	0.11	8.7	106.0	93.3	257.0	0.58	70.17	3.92	1.18
Median	71222.0	68.8	1.22	29.4	38.9	76681.0	0.47	12.1	252.0	105.0	535.0	1.35	85.38	9.24	2.18
Mean	66592.2	72.3	2.57	49.6	55.4	87693.3	0.86	14.6	372.8	108.6	853.7	3.83	78.18	17.99	5.13
P75	81480.0	84.0	2.71	38.6	89.2	120305.0	1.13	18.8	450.0	118.0	1192.0	4.56	94.12	26.26	6.78
P95	88671.4	115.6	10.30	88.6	139.0	173833.8	3.02	25.8	1280.4	134.0	2716.8	16.97	97.69	61.05	18.06
Max	92241.0	137.0	14.60	1022.0	210.0	209372.0	7.51	94.1	1426.0	360.0	3132.0	22.46	99.13	69.77	28.40
SD	17179.6	22.6	3.28	110.7	45.5	46655.1	1.15	10.4	366.2	32.7	807.9	5.35	20.13	19.21	6.31
MAD	13013.0	13.4	0.80	9.1	24.3	34692.0	0.39	4.2	152.2	13.0	350.0	1.18	9.88	6.87	1.57
IQR	28180.0	27.7	2.26	17.8	72.4	73193.0	1.02	10.1	344.0	24.7	935.0	3.98	23.94	22.35	5.60
CV (%)	25.8	31.3	127.48	223.4	82.1	53.2	133.78	71.1	98.2	30.2	94.6	139.78	25.75	106.80	123.15
CVR (%)	18.3	19.5	65.57	31.0	62.5	45.2	84.15	34.7	60.4	12.4	65.4	87.77	11.57	74.32	72.02
Distal zone (N = 32): 0-20 cm															
Min	29955.0	18.7	0.26	11.2	6.6	21893.0	0.006	4.2	47.3	61.8	94.8	0.01	8.61	0.03	0.06
P5	38741.6	22.0	0.30	13.1	6.9	24830.9	0.009	6.1	54.2	72.7	104.0	0.19	13.05	0.07	0.07
P25	55842.0	34.3	0.35	26.1	11.7	35437.5	0.04	8.1	65.1	86.5	143.5	0.40	38.44	0.35	0.17
Median	63351.5	41.9	0.41	31.8	20.4	38743.0	0.25	10.9	113.0	98.7	245.5	2.00	79.92	9.39	0.92
Mean	62368.7	41.6	0.46	39.8	26.4	41475.6	0.35	12.6	121.3	100.4	286.6	5.96	66.95	27.10	1.18
P75	70062.3	46.1	0.51	55.9	45.3	45840.0	0.65	15.9	149.3	105.0	394.3	5.87	94.23	54.74	1.99
P95	81736.6	65.2	0.77	73.3	54.7	61455.6	0.87	24.1	250.4	144.5	602.0	19.68	98.96	86.51	3.08
Max	87205.0	84.0	0.96	87.1	60.0	102294.0	1.10	28.0	285.0	200.0	713.0	55.38	99.46	91.11	5.36
SD	13679.7	14.1	0.17	20.4	17.8	14734.4	0.35	6.0	65.2	25.3	171.9	10.79	31.38	33.25	1.22
MAD	7541.5	6.1	0.09	15.1	12.8	6571.5	0.22	4.0	48.0	9.4	119.0	1.71	18.00	9.31	0.80
IQR	14220.3	11.7	0.16	29.8	33.5	10402.5	0.62	7.8	84.2	18.5	250.8	5.48	55.79	54.40	1.82
CV (%)	21.9	33.9	36.41	51.2	67.5	35.5	98.07	47.7	53.7	25.2	60.0	181.14	46.88	122.71	103.13
CVR (%)	11.9	14.4	20.99	47.5	62.7	17.0	90.67	36.2	42.5	9.5	48.5	85.52	22.52	99.11	86.89

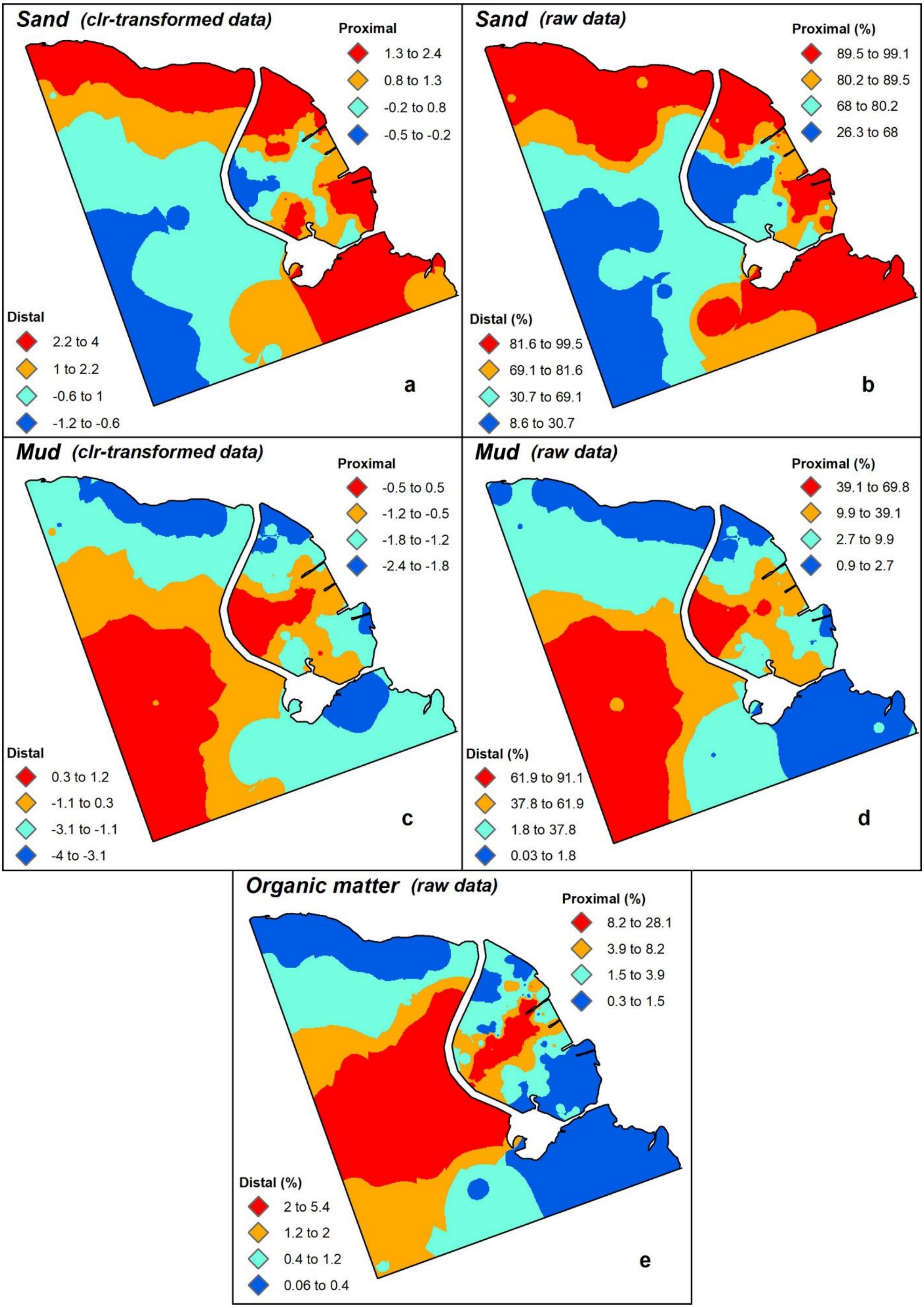


Fig. 8 Spatial distribution of sand, mud and organic matter produced by clr-transformed data (a and c) and raw data (b, d and e). Dealing with data closure does not significantly influence spatial patterns of the grain size fractions.

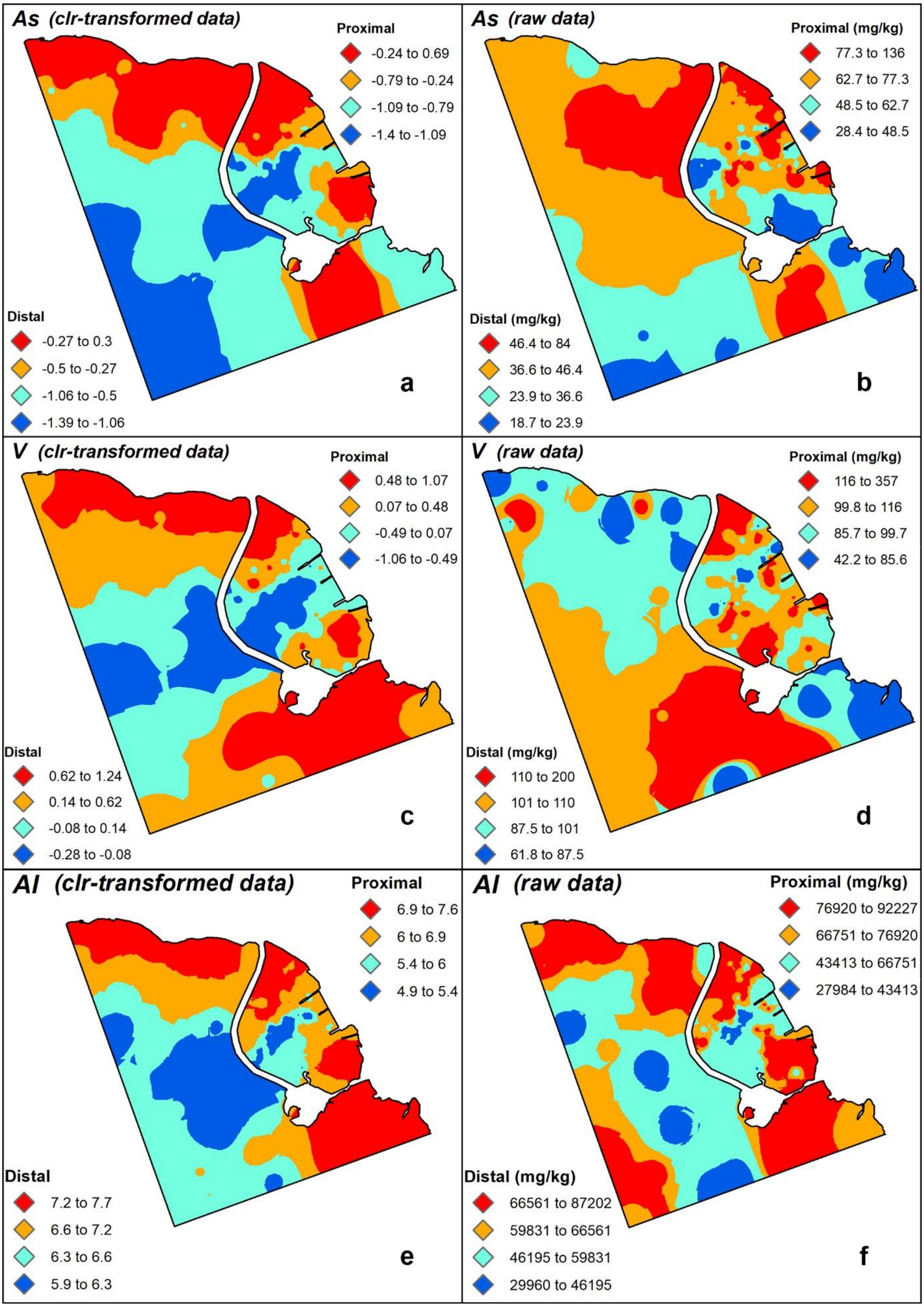


Fig. 9 Arsenic, vanadium, aluminum, iron, chromium, nickel, mercury, cadmium, copper, lead and zinc geochemical maps produced by clr-transformed data (a, c, e, g, i, k, m, o, q, s and u, respectively) and raw data (b, d, f, h, j, l, n, p, r, t, and v, respectively).

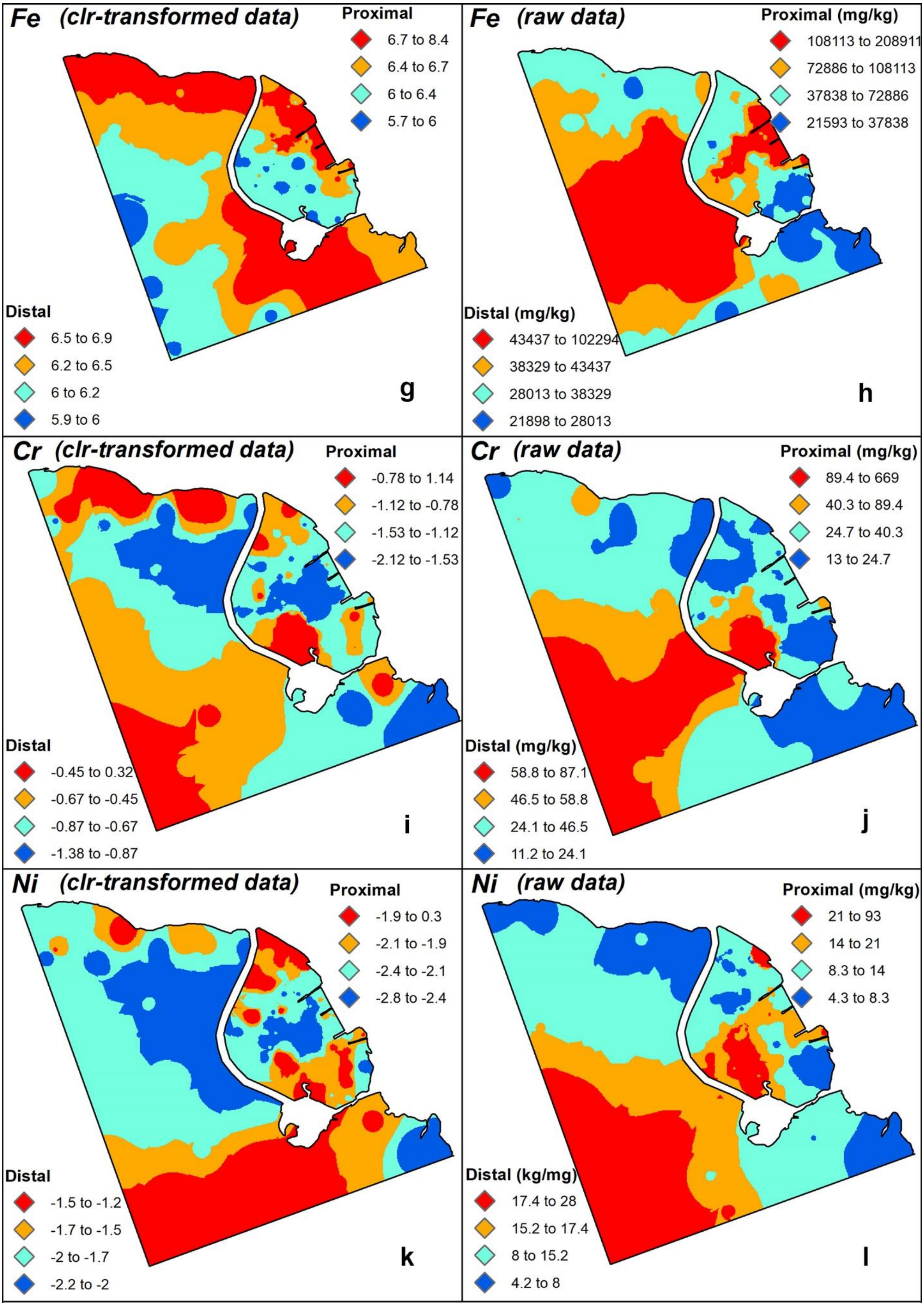


Fig. 9 (continued)

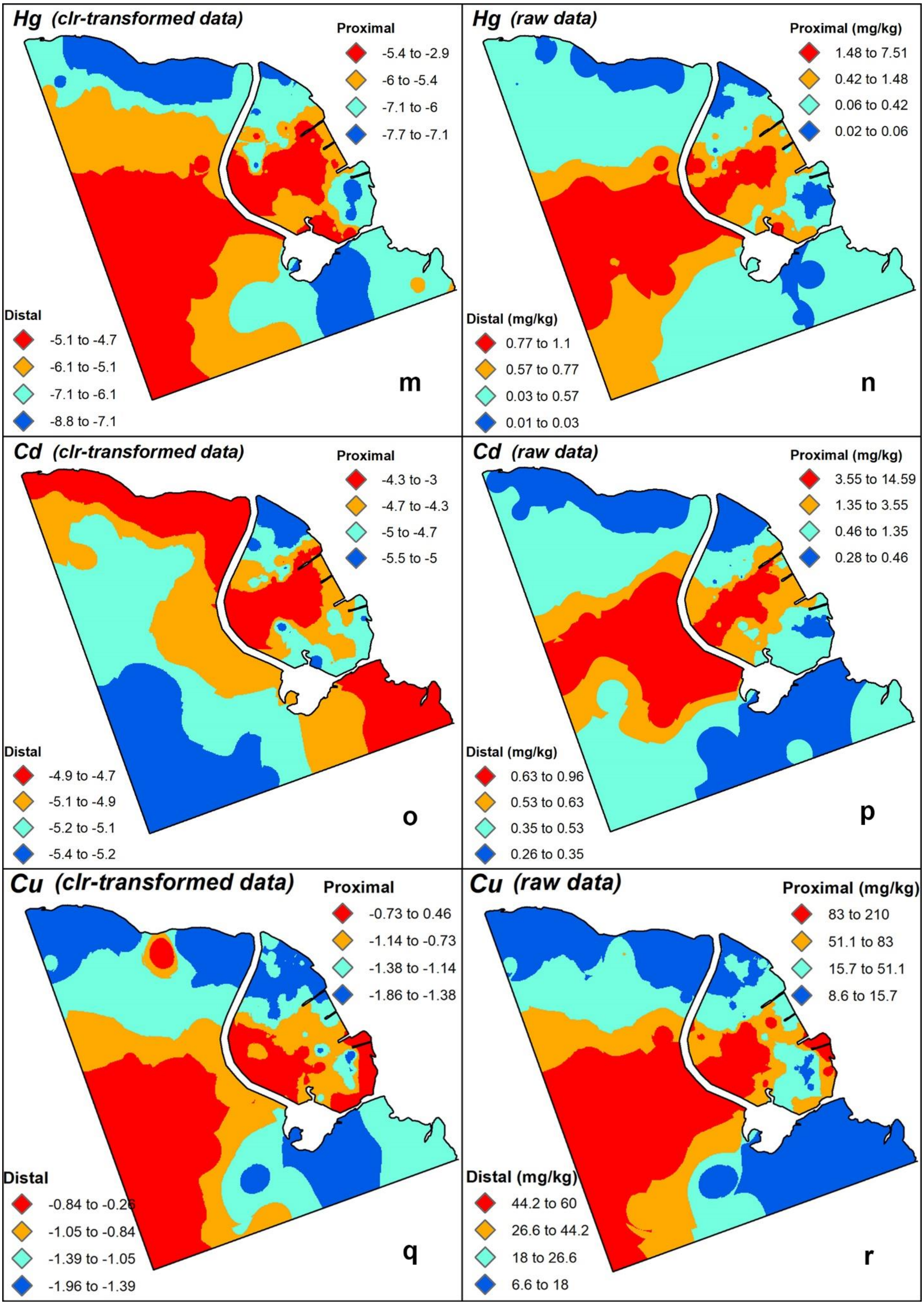


Fig. 9 (continued)

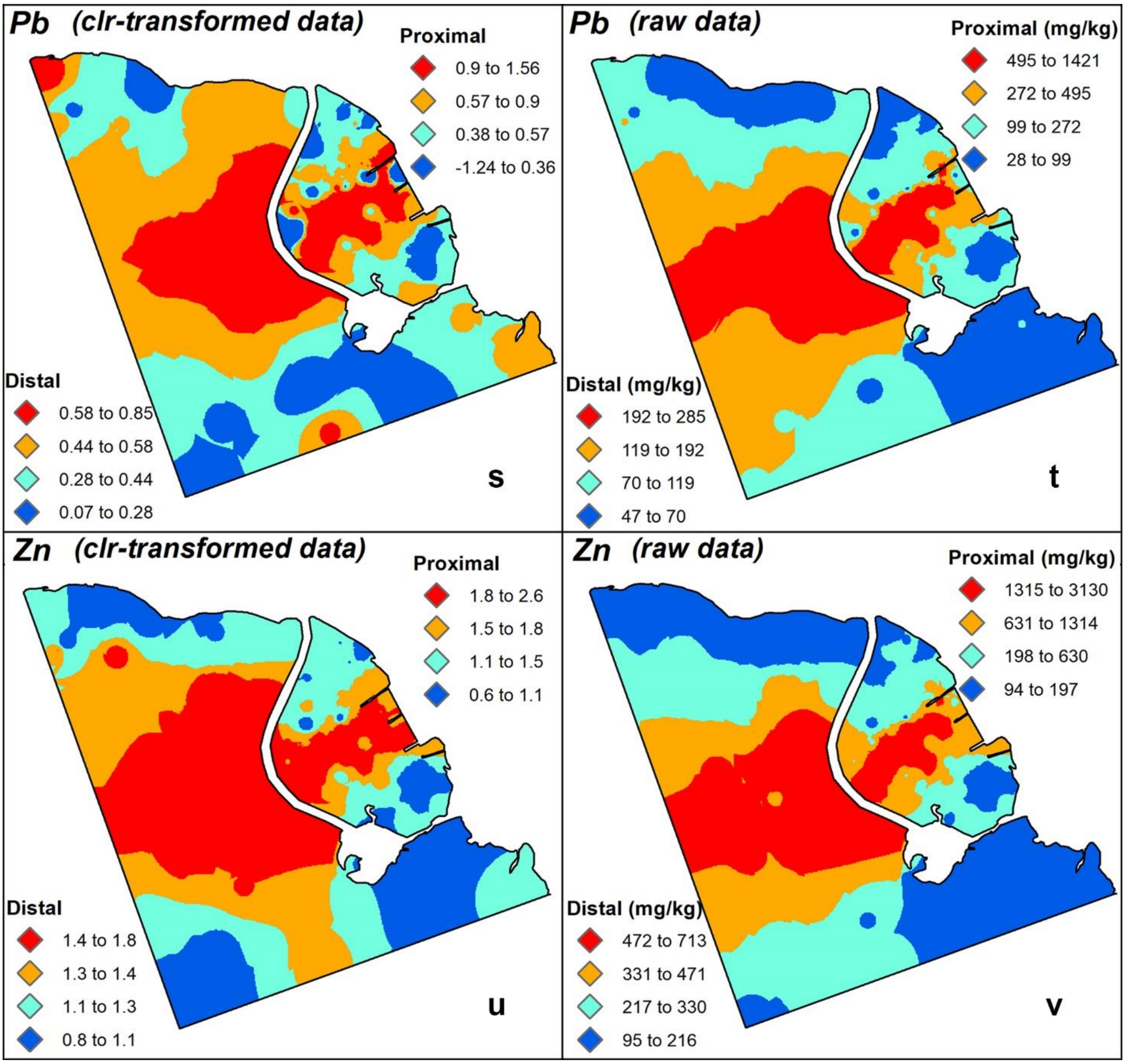


Fig. 9 (continued)

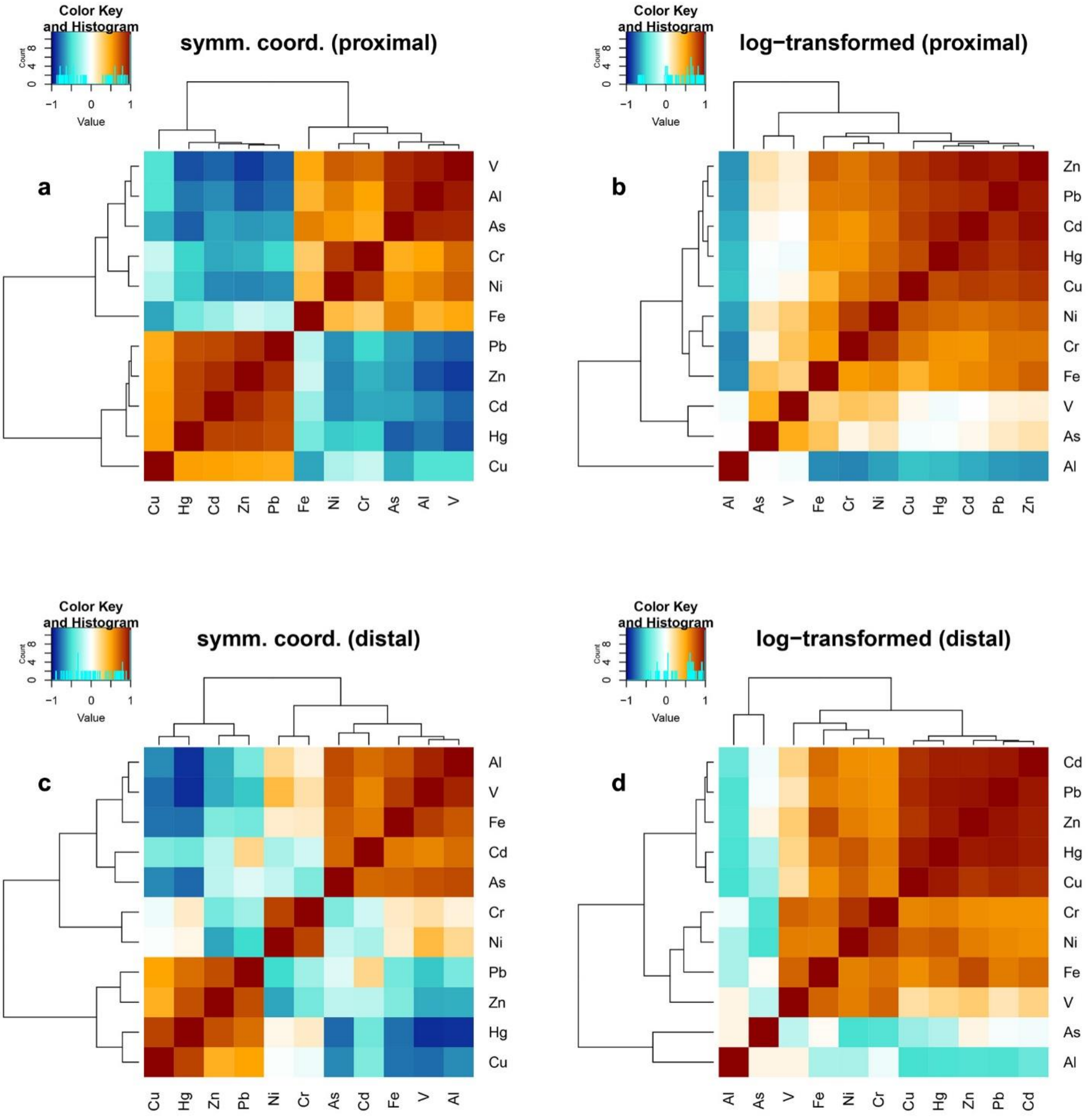


Fig. 10 The heatmaps of Spearman correlation between the measured chemicals in proximal (a and b) and distal (c and d) sediments. The left and right panels refer to the correlation analysis applied to the symmetric coordinates and log-transformed data, respectively.

4.1.4. The processes governing geochemistry of the sea sediments

Regarding the results of correlation analysis (Fig. 10), CoDA was integrated with PCA to find the main geochemical processes. The rise of hot thermal fluids through underwater fumaroles (Albanese et al., 2010), minerals, industrial activity (Armiento et al., 2020) together with discharge of As-rich groundwater (Aiuppa et al., 2006; De Vivo et al., 2008) and a mixture of thermal fluids and municipal wastewaters (Giglioli et al., 2020) in the sea were introduced as the probable sources of arsenic in the study area. In addition, the fine fraction of sediments and organic materials are effective metal scavengers (Keshavarzi et al., 2015; Sciarrillo et al., 2020). Therefore, sand and mud were treated as the second set of compositional data whilst organic matter (OM) along with the shortest distance of each sample to fumaroles, coastline and seepages (i.e.

dist_fumarole, dist_coast and dist_seepage, respectively) were considered as external non-compositional variables in PCA. Presence of data outliers was verified prior to statistical analysis via scatterplots of the symmetric coordinates of elements (Fig. 11). Hence, the classical and robust PCA were carried out to evaluate impact of the outliers on the outcomes and it was apparent that the robust method had a better performance because it explained a greater variance in the dataset (Figs. 12 and 13). Regarding the proximal sediments, biplot of the first two principal components which accounted for 89.35% of the total variance was very informative and grouped the variables in two associations (Fig. 12a): (1) Hg, Cd, Cu, Pb, Zn, OM, mud, dist_coast and dist_seepage; and (2) Al, As, V, Fe, Cr, Ni, sand and dist_fumarole. On the other hand, 94.73% of variance was explained by PC1 and PC2 in the distal sediments (Fig. 12b). Compared with the PZN, shorter rays of the elements highlight their lower variation in the sediments deposited far from the Bagnoli industrial site. According to the PCA biplot (Fig. 12b), two relevant elemental associations are as follows: (1) Hg, Cu, Zn, Pb, OM, mud, dist_seepage, dist_fumarole and dist_coast; and (2) Al, As, V, Fe, Ni, Cr and sand. Cadmium was excluded because of the lowest variability in PCA. This might suggest a non-industrial source for Cd (e.g., from seepages), but further investigation is needed to confirm this assumption.

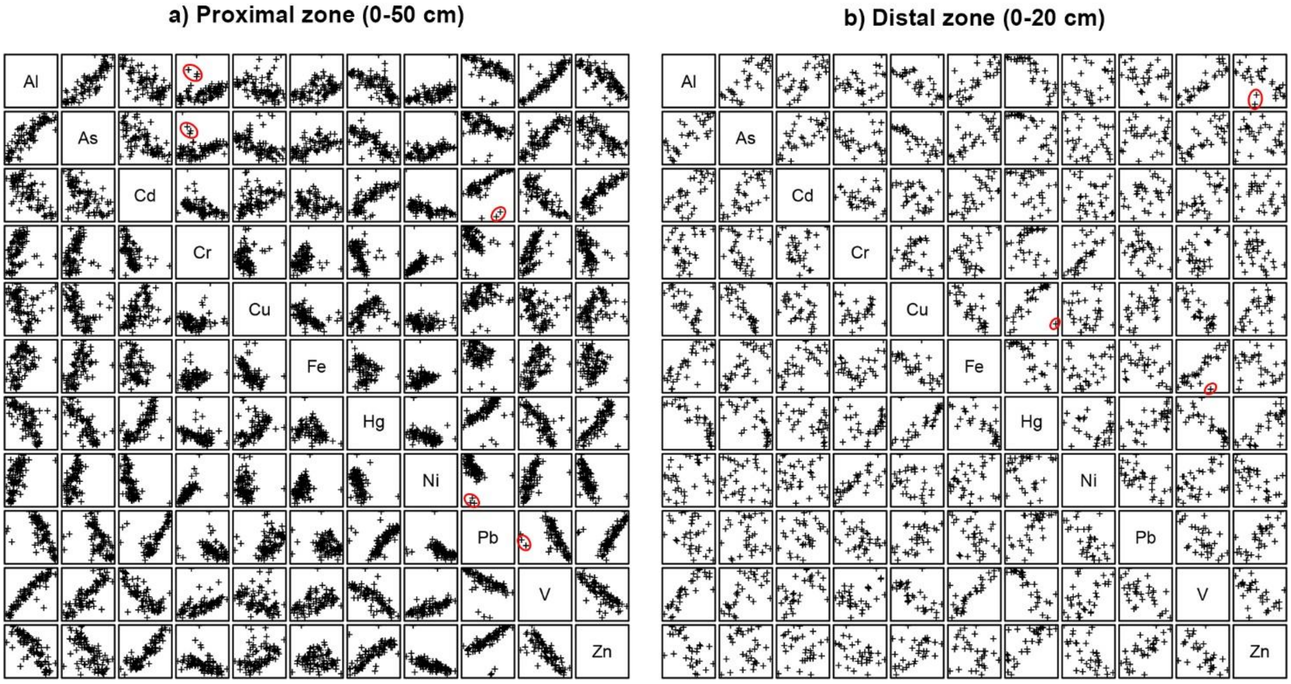


Fig. 11 Scatterplot matrices of the elements using symmetric pivot coordinates: (a) proximal zone and (b) distal zone. Some of the possible data outliers are illustrated with red circles.

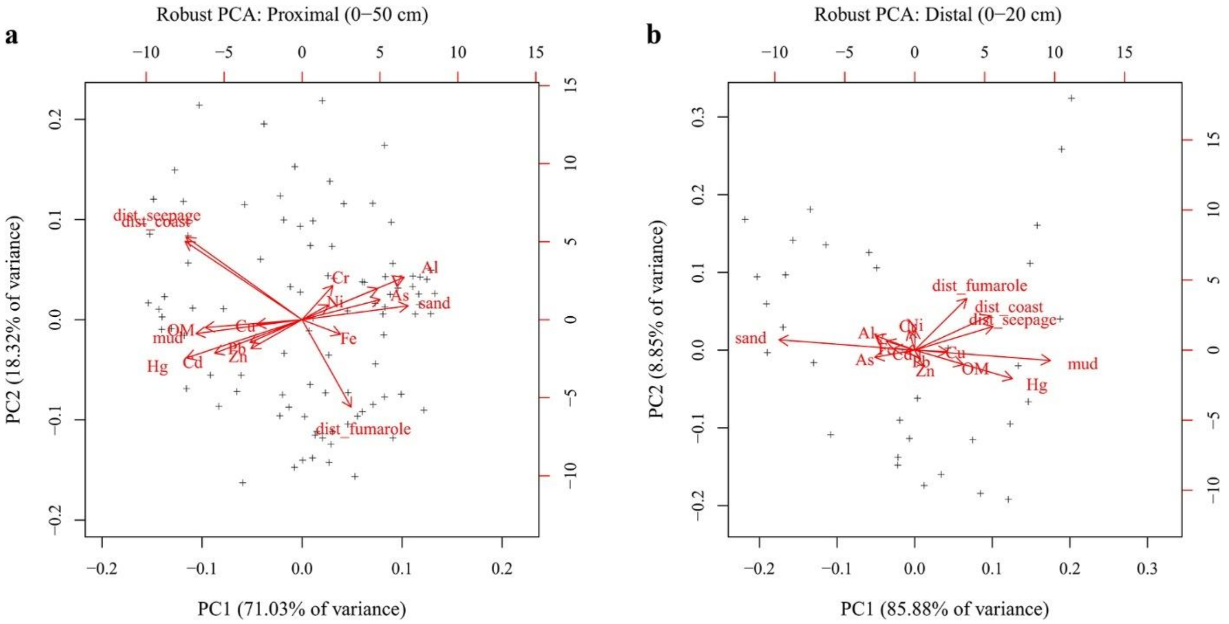


Fig. 12 Robust principal component analysis of the proximal and distal (a and b, respectively) datasets. Amongst the variables, the granulometry and element contents are two different compositions, and the rest are non-compositional variables. The distance from coast, seepage and fumarole was represented by dist_coast, dist_seepage and dist_fumarole, respectively.

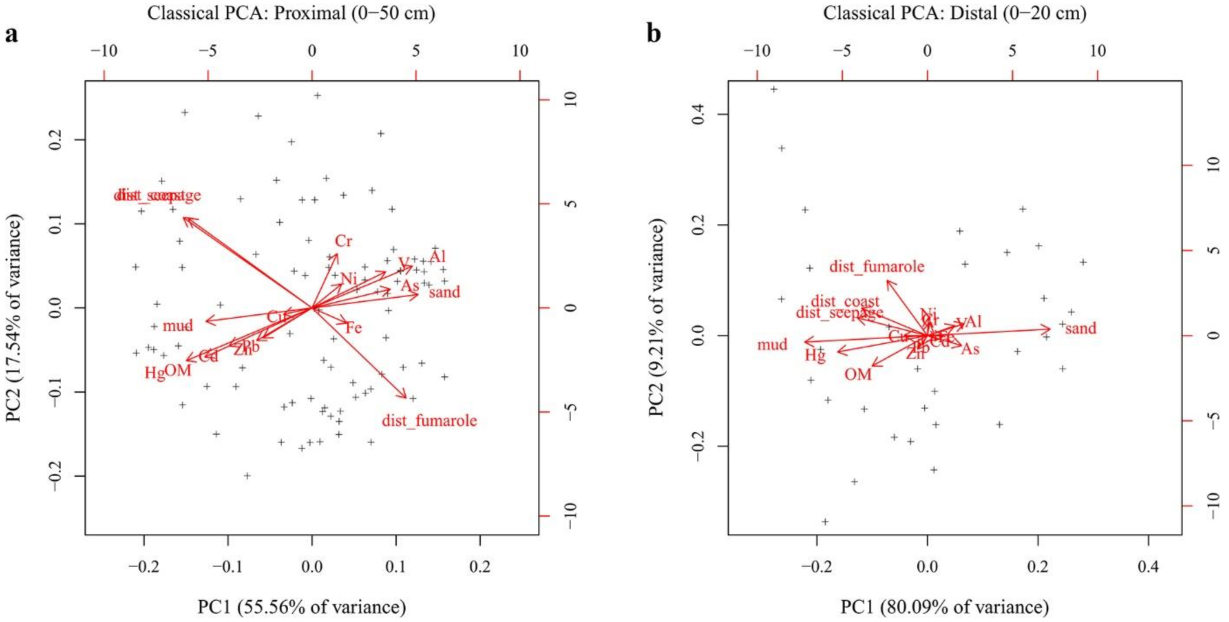


Fig. 13 Classical principal component analysis of the proximal and distal (a and b, respectively) datasets. Amongst the variables, the granulometry and element contents are two different compositions, and the rest are non-compositional variables. The distance from coast, seepage and fumarole was represented by dist_coast, dist_seepage and dist_fumarole, respectively.

The elemental association of Hg, Cd, Cu, Pb and Zn in the PZN is probably related to the coastal landfill. Occurrence of OM and the fine fraction of sediments with these PTEs represents their tendency to adsorb metals in aqueous environment, being in agreement with Keshavarzi et al. (2015) and Sciarrillo et al. (2020). Moreover, the strong correlation between total polycyclic aromatic hydrocarbons (PAHs) and total organic carbon (Arienzo et al., 2017) together with similar spatial distribution of PAHs, Pb and Zn (Romano et al., 2009) in the study area support attributing this principal component to anthropogenic activity. Regarding Fig. 12a, there is a positive correlation between Hg, Cd, Cu, Pb, Zn and the distance to coastline and seepages. It might occur because of wave turbulence at coast (Arienzo et al., 2019), higher mud and OM contents far from the shoreline and insignificant impact of the seepages on the PTEs concentrations. This group of chemicals are uncorrelated with the distance to fumaroles which highlights their anthropogenic source. The similar statistical relationships between the abovementioned variables in the DZN (Fig. 12b) feasibly indicate that the sediment pollution is extended to the deeper parts of Pozzuoli Bay. Increasing distance from fumaroles leads to higher levels of Hg and Cu in distal sediments, representing another piece of evidence for their non-geogenic origin.

The previous investigations in soil media showed that the association of Al, As and V can be related to the pyroclastic deposits of Phlegrean Fields (Thiombane et al., 2019). Significant variance and correlation of these variables in PCA may indicate that As is incorporated in the crystalline structure of the sand fraction in bed sediments of PZN. Furthermore, although arsenic is uncorrelated with distance from fumaroles, it is negatively correlated with distance from coastline and seepages (Fig. 12a). Thus, increasing the distance from shoreline and seepages results in decreasing As concentration in the sediments. The present dataset does not show the relative importance of submarine arsenic-rich groundwater discharge (Aiuppa et al., 2006; De Vivo et al., 2008) in each zone. If it affects PZN and DZN equally, it can be inferred that surface discharge of mixed thermal fluids and municipal wastewaters (e.g., the Conca d’Agnano seepage in Figs. 1 and 2) is likely the key mechanisms for transporting As to PZN. It is worth mentioning that arsenic levels in PZN are higher than DZN (Table 3 and Fig. 9b). The correlation between As, Fe and sand (Fig. 12a) may indicate that sand fraction is a potential arsenic carrier in the study area. The naturally available iron-rich sand can be a promising As adsorbent in seawater pH (Panthi and Wareham, 2011). Armiento et al. (2020) also introduced Fe oxyhydroxides as a controlling factor for As distribution in the deep sediment layers of east Pozzuoli Bay. High dependence of Ni and Fe to the sandy sediments in the proximal area is probably derived from the anthropogenic sands composed of tar and oil residues, blast furnace, slag and iron pellets (Albanese et al., 2010). In the DZN, low contribution of Al, As, V, Fe, Ni and Cr in PCA (Fig. 12b) might explain the important role of discharging thermal fluids through PZN seepages (Giglioli et al., 2020) in governing sediment geochemistry.

In order to better interpret the results of principal component analysis, the interpolated PC1 scores were represented in Fig. 14. In sediments of the proximal zone, the negative PC1 scores (i.e. the dependence of Hg, Cd, Cu, Zn and Pb to the fine fraction and OM) coincide with the Bagnoli industrial site. Presence of these polluted sediments in the offshore sector of DZN can also be verified by the Hg–Cu–mud–OM association. On the other hand, Al, As, V and Fe (together with Ni and Cr in PZN) form an association along the shoreline which might be regulated by mobilization of the fine and very fine sand that chemically interacted with seepages (Fig. 14). Hence, regarding the spatial distribution of PC1 scores and sediment dispersal towards the open sea, geochemical characteristics of the sediments in DZN likely depends on the sediments transported from PZN.

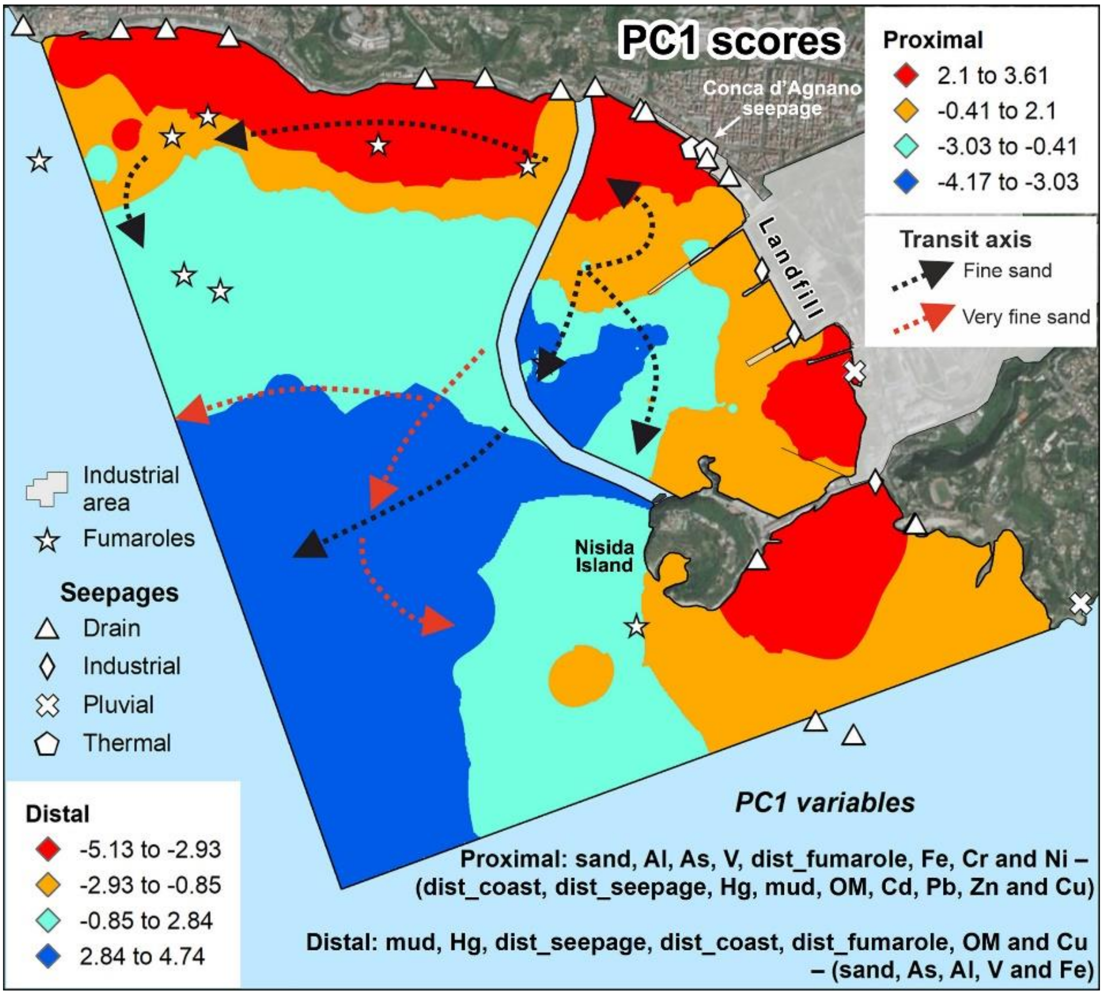


Fig. 14 PC1 score map of the sediments collected from the proximal (0-50 cm) and distal (0-20 cm) zones. Due to low contribution of Cd in robust PCA and very low correlation of Ni, Cr, Pb and Zn with PC1, they were excluded from PC1 of the DZN. The transit axes of fine and very fine sand (Arienzo et al., 2017) are also illustrated.

4.1.5. Sediment pollution

Considering the identified geochemical associations in robust PC1 (Fig. 12), the variables of each compositional dataset were divided in two groups (i.e. the numerators and denominators of Eqs. (8) and (9)) and the following balances were calculated for PZN according to Eq. (7):

$$\tilde{z}_{\text{granulumetry}} = \sqrt{\frac{1}{2}}\ln\frac{\text{sand}}{\text{mud}} \tag{8}$$

$$\tilde{z}_{\text{elements}} = \sqrt{\frac{30}{11}}\ln\frac{(\text{Al}\times\text{As}\times\text{Fe}\times\text{Ni}\times\text{V}\times\text{Cr})^{\frac{1}{6}}}{(\text{Hg}\times\text{Cd}\times\text{Cu}\times\text{Pb}\times\text{Zn})^{\frac{1}{5}}} \tag{9}$$

Likewise, the computed balances for the DZN are as follows:

$$\tilde{z}_{\text{granulumetry}} = \sqrt{\frac{1}{2}}\ln\frac{\text{sand}}{\text{mud}} \tag{10}$$

$$\tilde{z}_{\text{elements}} = \sqrt{\frac{8}{6}} \ln \frac{(\text{Al}\times\text{As}\times\text{Fe}\times\text{V})^{\frac{1}{4}}}{(\text{Hg}\times\text{Cu})^{\frac{1}{2}}} \tag{11}$$

In each zone, linear regression model was then applied to the $\tilde{z}_{\text{granulumetry}}$ as independent variable and $\tilde{z}_{\text{elements}}$ as dependent variable, and the obtained standardized residuals were mapped via MIDW approach as described in the section 3.2.1. As Fig. 15 shows, distribution of the positive values mostly represents the nonindustrial source (i.e. pyroclastic deposits, discharge of mixed thermal fluids and municipal wastewaters, and arsenic-rich groundwater discharge) of chemicals in the sand fraction, particularly along the coastline. The high values towards the open sea likely show elevated Cr and Ni concentrations in the fine fraction (Wang et al., 2015). However, the samples with negative standardized residuals occur in front of Bagnoli industrial site which can also be traced in the DZN. The negative values in south and east of DZN might be attributed to the seepages, especially the industrial seepage. Wave-induced sediment transport probably disturbed the spatial patterns.

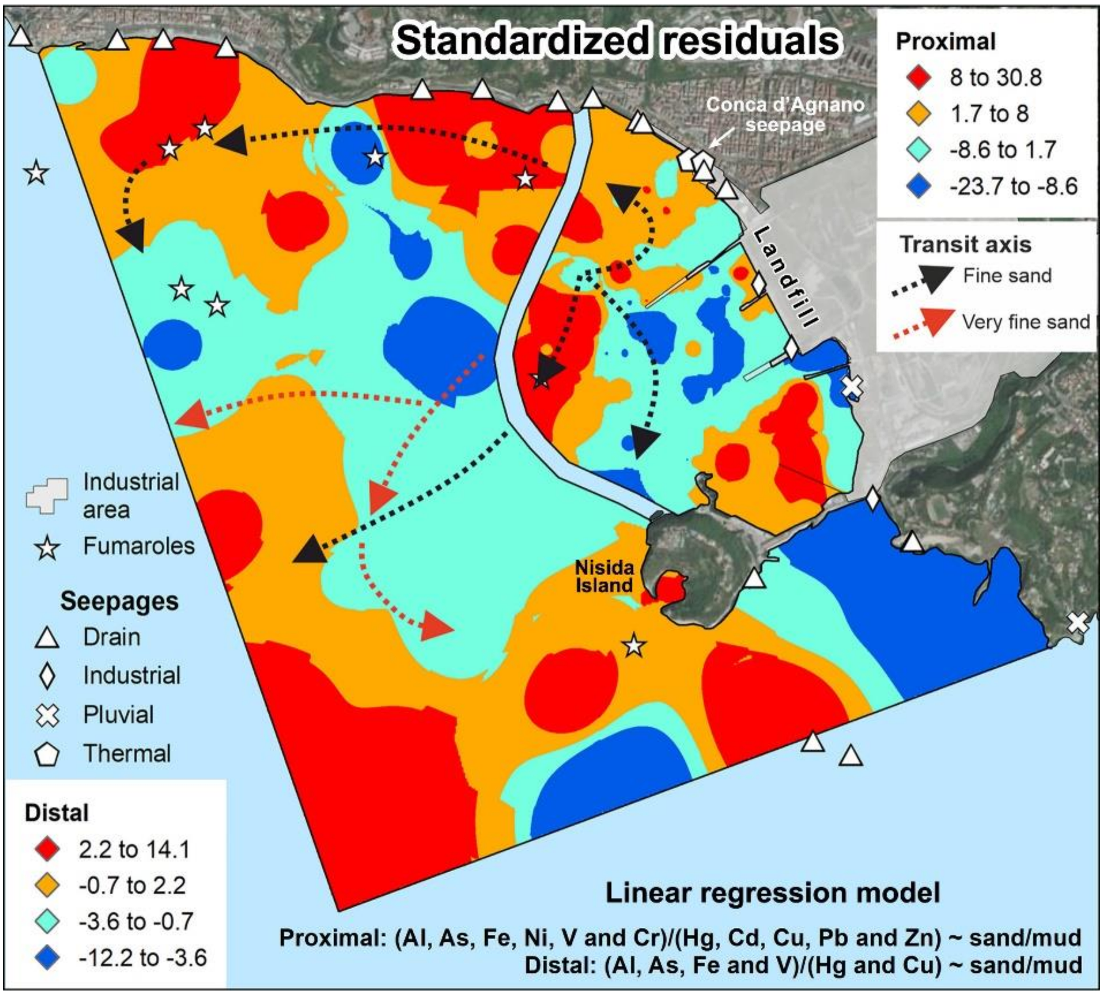


Fig. 15 Spatial distribution of the standardized residuals in the proximal and distal zones. The transit axes of fine and very fine sand (Arienzo et al., 2017) are depicted as well.

Table 4 Statistical summary of the measured variables in 50-100 cm, 100-150 cm, 150-200 cm and 200-300 cm depth intervals of the sediment cores collected in the proximal zone. The elements are in mg/kg and the rest are in percentage.

	Al	**As**	**Cd**	**Cr**	**Cu**	**Fe**	**Hg**	**Ni**	**Pb**	**V**	**Zn**	**Gravel**	**Sand**	**Mud**	**OM**
50-100 cm (N = 58)															
Min	47686.0	35.3	0.27	8.6	6.8	30159.0	0.01	4.9	16.1	78.6	83.0	0.00	19.37	0.92	0.17
P5	48954.3	43.9	0.32	10.8	8.2	31484.7	0.02	5.8	33.7	81.9	96.5	0.01	30.58	2.00	0.21
P25	64732.8	56.8	0.39	17.1	11.7	39205.3	0.08	7.0	63.1	90.3	170.3	0.31	55.75	4.60	0.36
Median	79212.0	65.9	0.80	24.8	19.5	51032.5	0.27	9.2	132.5	104.5	366.5	0.95	88.37	7.86	1.83
Mean	76270.9	73.4	3.81	31.8	54.4	65001.2	1.10	12.1	405.0	106.9	973.6	2.71	75.24	22.06	3.78
P75	89709.5	80.5	3.33	34.0	85.0	81800.0	1.72	14.9	498.3	117.0	1152.8	3.26	94.41	38.26	6.18
P95	93729.9	115.6	17.28	72.4	185.1	143150.6	3.69	23.9	1550.6	144.2	3306.4	10.66	96.47	66.20	13.16
Max	94939.0	206.0	26.10	142.8	204.0	158329.0	6.56	42.7	2486.0	180.0	5360.0	12.22	98.60	79.68	16.84
SD	15332.0	30.0	6.15	25.3	59.1	35377.5	1.52	7.3	556.0	21.2	1230.5	3.48	23.63	22.62	4.46
MAD	11720.5	10.4	0.49	9.2	11.2	15377.5	0.25	3.0	95.4	14.0	266.5	0.86	8.14	6.03	1.55
IQR	24976.8	23.7	2.94	16.9	73.3	42594.8	1.64	7.9	435.2	26.8	982.5	2.95	38.66	33.66	5.82
CV (%)	20.1	40.9	161.32	79.4	108.6	54.4	138.11	60.5	137.3	19.8	126.4	128.48	31.41	102.57	117.72
CVR (%)	14.8	15.7	60.63	37.1	57.3	30.1	93.20	33.2	72.0	13.4	72.7	90.10	9.21	76.70	84.93
100-150 cm (N = 55)															
Min	36109.0	30.1	0.25	8.6	6.0	29186.0	0.001	4.0	20.0	76.4	76.5	0.00	32.72	1.63	0.14
P5	61731.6	34.3	0.27	9.6	6.2	31845.7	0.01	4.9	24.7	80.5	82.2	0.09	42.34	1.82	0.18
P25	81057.0	54.3	0.33	13.1	7.8	35106.0	0.02	6.0	35.9	91.2	96.7	0.29	78.59	3.03	0.26
Median	88784.0	65.6	0.38	17.9	9.6	38485.0	0.04	7.8	51.2	102.0	108.0	1.46	89.31	4.69	0.54
Mean	85065.0	66.5	1.40	21.4	27.4	47741.3	0.52	9.2	192.3	103.5	516.0	2.56	82.64	14.81	2.03
P75	93341.5	77.3	0.65	25.3	19.8	45558.5	0.17	10.2	152.0	114.5	296.0	3.56	95.15	19.63	2.27
P95	95975.4	95.6	6.83	44.5	129.2	91046.2	3.33	16.2	1148.9	132.0	2587.5	9.59	95.96	56.48	8.44
Max	102100.0	138.0	15.70	84.4	152.0	164867.0	4.75	46.0	1674.0	148.0	4926.0	12.41	97.27	65.70	19.08
SD	12557.4	21.3	2.75	13.2	38.5	26812.6	1.12	6.1	357.9	16.1	975.8	2.99	17.34	17.88	3.51
MAD	5127.0	11.7	0.10	5.4	3.4	4249.0	0.04	2.0	22.7	11.0	24.0	1.31	6.01	2.64	0.32
IQR	12284.5	23.0	0.32	12.2	12.0	10452.5	0.15	4.2	116.2	23.3	199.4	3.27	16.56	16.60	2.01
CV (%)	14.8	32.1	196.82	61.7	140.7	56.2	214.87	66.0	186.1	15.5	189.1	117.03	20.99	120.79	172.86
CVR (%)	5.8	17.8	26.32	30.2	35.3	11.0	86.36	25.5	44.3	10.8	22.2	89.44	6.73	56.37	59.26

Table 4 (continued)

	Al	**As**	**Cd**	**Cr**	**Cu**	**Fe**	**Hg**	**Ni**	**Pb**	**V**	**Zn**	**Gravel**	**Sand**	**Mud**	**OM**
150-200 cm (N = 50)															
Min	27725.0	27.8	0.24	8.0	6.2	26203.0	0.001	4.3	18.2	79.9	79.7	0.00	41.23	1.36	0.16
P5	52469.9	32.9	0.27	9.8	6.5	29269.9	0.005	4.9	20.8	84.3	82.1	0.10	44.15	2.19	0.17
P25	86031.8	55.9	0.31	12.5	7.1	34742.0	0.008	5.7	29.3	95.0	90.5	0.38	80.96	3.82	0.22
Median	89779.0	67.9	0.35	14.9	8.6	37305.5	0.02	7.4	39.6	103.0	98.8	0.66	93.17	4.84	0.29
Mean	85882.0	67.1	1.05	21.9	20.8	40599.7	0.35	8.4	167.3	106.3	481.4	1.55	84.96	13.49	1.33
P75	93575.0	80.1	0.42	20.7	12.3	40245.5	0.07	9.6	57.2	118.0	132.0	1.56	94.75	18.59	0.76
P95	97332.1	107.7	5.14	41.4	106.5	68409.1	2.65	18.0	879.8	136.7	2618.3	5.27	96.98	54.36	8.42
Max	101340.0	114.0	11.40	216.0	163.0	114450.0	4.71	22.2	2481.0	140.0	8535.0	14.23	97.94	58.64	12.94
SD	14393.7	21.5	2.22	29.5	33.6	15160.4	0.94	3.9	430.9	16.3	1366.4	2.37	15.66	15.76	2.85
MAD	3908.0	12.2	0.05	3.7	1.9	2883.5	0.01	1.8	12.3	11.4	10.7	0.46	3.41	2.38	0.11
IQR	7543.3	24.2	0.11	8.2	5.2	5503.5	0.06	3.9	27.9	23.1	41.5	1.18	13.79	14.77	0.54
CV (%)	16.8	32.0	211.89	134.9	161.9	37.3	268.40	47.0	257.6	15.3	283.8	153.38	18.43	116.88	214.50
CVR (%)	4.4	17.9	14.29	24.8	21.9	7.7	70.27	24.0	31.1	11.0	10.8	70.27	3.66	49.17	37.93
200-300 cm (N = 14)															
Min	73847.0	48.9	0.24	15.7	5.40	30816.0	0.001	5.96	23.7	86.8	86.5	0.00	59.06	2.24	0.13
P5	74746.6	50.5	0.26	15.8	6.12	33835.3	0.002	5.99	23.8	88.8	87.7	0.08	75.03	2.52	0.14
P25	88828.3	62.3	0.30	17.6	7.85	40329.5	0.006	6.84	30.1	123.0	91.7	0.25	91.53	3.64	0.22
Median	94010.0	88.2	0.37	23.1	8.31	41789.5	0.014	7.79	34.7	126.5	105.5	0.93	93.15	4.88	0.33
Mean	98076.7	94.8	0.61	22.0	16.82	49279.9	0.16	8.01	83.8	129.7	223.3	1.33	90.45	8.22	0.69
P75	96423.5	110.4	0.44	25.2	14.60	45658.0	0.04	9.31	61.5	141.5	121.8	1.97	94.44	7.78	0.42
P95	133816.5	167.3	1.77	28.3	54.57	85085.2	0.82	10.41	296.4	164.5	807.9	3.76	96.36	24.67	2.63
Max	201371.0	229.0	2.52	28.6	76.80	95530.0	1.38	10.80	429.0	169.0	971.0	4.65	97.33	40.38	4.48
SD	30751.8	47.0	0.63	4.6	19.67	18552.3	0.38	1.59	113.7	24.5	274.8	1.40	9.58	9.91	1.15
MAD	2861.0	26.1	0.08	4.1	1.63	3214.5	0.01	1.23	8.1	13.5	15.9	0.73	1.55	1.78	0.11
IQR	7595.3	48.1	0.14	7.6	6.76	5328.5	0.04	2.47	31.4	18.5	30.1	1.72	2.90	4.14	0.20
CV (%)	31.4	49.5	104.04	20.8	116.93	37.7	237.37	19.84	135.6	18.9	123.1	105.40	10.59	120.54	168.15
CVR (%)	3.0	29.6	21.62	17.6	19.61	7.7	60.71	15.73	23.4	10.7	15.1	78.78	1.67	36.50	33.33

4.2. Sediment geochemistry along the cores

Considering the sediments from different depths (Tables 3 and 4) in PZN, the average OM contents and CVR values are higher at surface (<100 cm depth). The same vertical pattern was also observed for mud, Cu, Pb, Zn, Cd, Hg, Cr, Ni and Fe. On the other hand, variation of sand, Al, As and V are almost similar. In the next step, the sediment cores that reach 300 cm depth (the bottom right panel of Fig. 16) were selected and the average values for the analyzed elements, grain size fractions and OM in each depth interval were plotted in Fig. 16. Furthermore, the clr-transformed values of the metal(loid)s and granulometry data were included in each graph. It is evident that opening the data changes the vertical trend of some parameters (e.g., Ni and Cr) and better illustrates variation of As and V with depth. Regarding the clr-transformed values (except for OM), two groups of variables can be identified: (1) Al, As, Cr, V, Ni and sand; and (2) Cu, Cd, Hg, Pb, Zn, mud and OM. The minerals such as feldspar (plagioclase and potassium feldspar), biotite and pyroxene in the sand fraction (Romano et al., 2018) can be source of the elements in the first group. It seems that Fe shows an increasing trend towards depth, but high Fe contents at surface might be due to the iron pellets in the sea sediments (Albanese et al., 2010). In addition, the high proportion of haematite in the 0-50 cm depth interval of the sediment cores collected between the piers (Molisso et al., 2020) and the coarse fragments of coal and blast furnace slag in the upper 120 cm layer of sediments (Romano et al., 2018) can lead to the irregular Fe profile. Vertical profiles of the elements in the second group likely suggest that although industrial activity was stopped several decades ago, Bagnoli landfill is still a pollution source. Ignoring compositional nature of the geochemical data avoids identifying some of the observed patterns in Fig. 16.

Because clr-transformation is sub-compositionally incoherent, PCA was also performed for dimension reduction and for verifying the abovementioned visual inspections. The variation retained by the first two principal components of robust PCA is approximately 6% greater than that of classical PCA (93.43% in Fig. 17 vs. 87.19% in Fig. 18). Therefore, biplot of robust PC1 and PC2 is more informative. Two associations of parameters can be determined in Fig. 17a: (1) Hg, Cd, Cu, Pb, Zn, mud and OM; and (2) Al, As, V, Fe, Ni, Cr and sand. This classification of elements is consistent with the observed trends in Fig. 16. There is a strong association between Hg, Cd, Cu, Pb, Zn, mud and OM at <100 cm depth. These PTEs are anthropogenic pollutants as it was discussed previously. However, Al, As, V, Fe, Ni, Cr and sand are highly correlated in deeper parts (100-300 cm) which probably represents their geogenic source. Armiento et al. (2020) attributed the elevated arsenic contents in the sediments from 200 to 300 cm below surface to iron oxyhydroxides. The same groups of variables in PCA of the sediments from 0 to 50 cm depth (Fig. 12a) and the interpolated map of PC1 scores (Fig. 14) indicates extent of the polluted sediments in front of Bagnoli coastal landfill. According to Fig. 17, thickness of this polluted layer is about 100 cm along the shoreline. Sharp and Nardi (1987) estimated a sedimentation rate of 0.4 cm/y in the study area and concluded that maximum pollution occurred approximately 70 years ago. Because the current sampling campaign was conducted in 2017, it can be inferred that the most polluted sediment layer is about 40 cm below surface. However, some mechanisms such as sediment burial (Daliri et al., 2020) and wave turbulence (Arienzo et al., 2019) might affect vertical distribution of chemicals at <100 cm depth. Diagenesis is another process that might alter chemical phases and geochemical behavior of elements, but further investigation is required to confirm this assumption. In the form biplot (Fig. 17b) which favors visualization of individuals, the existing trend from natural samples at depth to contaminated samples at surface depicts pollution history of the study area.

The discussion about subsurface geochemistry of marine sediments can be informative in the present study because the profiles of variables (e.g., Cu and Fe in Figs. 19 and 20, respectively) generally agree with the high-resolution measurements along isolated cores (Damiani et al., 1987; Romano et al., 2018). Nevertheless, the site-dependent geochemical signatures (e.g., Figs. 17, 19 and 20) in front of the industrial plants highlight complexity of the sedimentary environment and importance of number and length of sediment cores together with sampling resolution along each core.

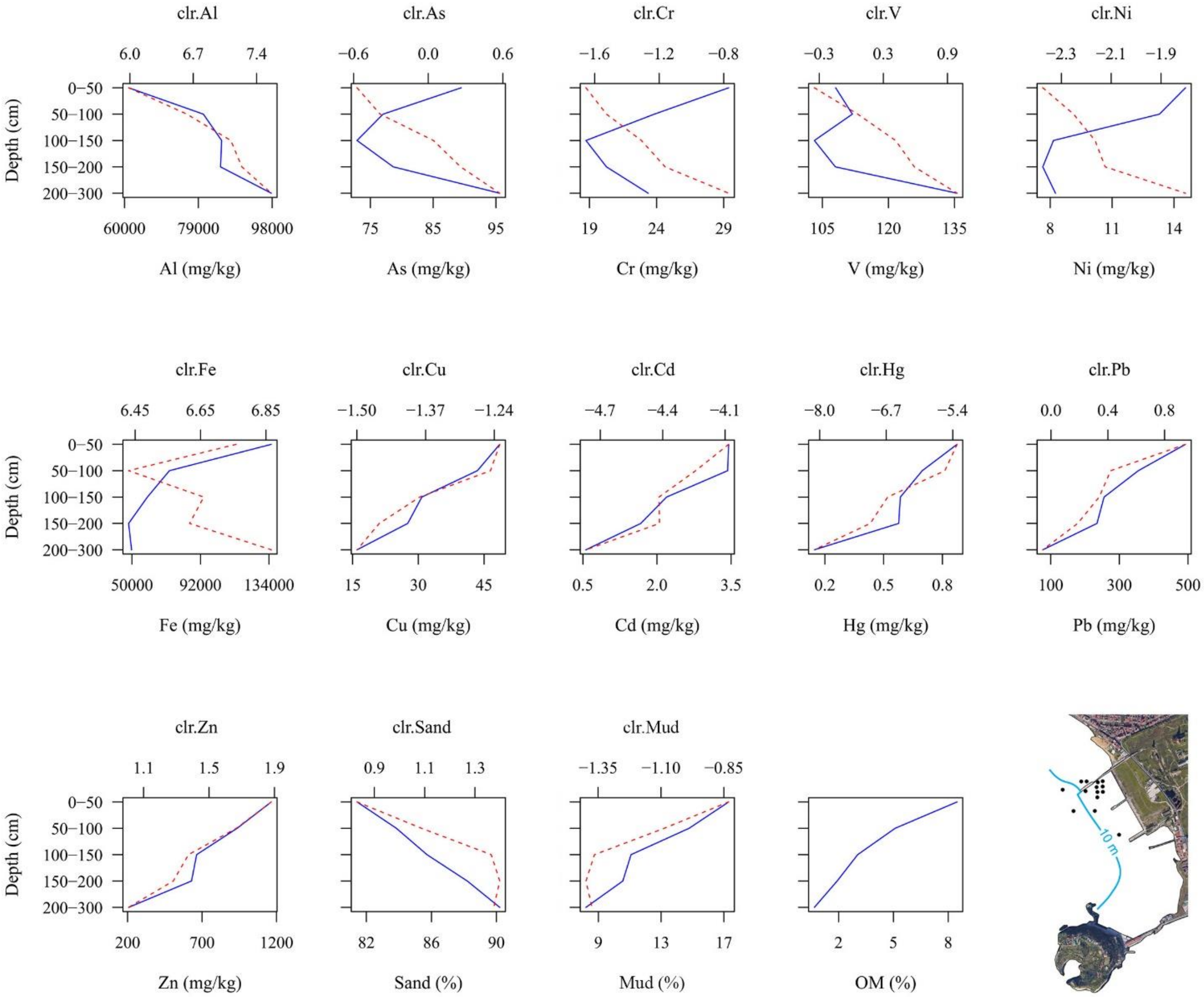


Fig. 16 Variation of the chemical elements, granulometry and organic matter (OM) content of the sediments from 0 to 300 cm depth. The solid blue line represents the measured values and the dashed red line shows the clr-transformed data. Location of the selected core samples are indicated in the bottom right map.

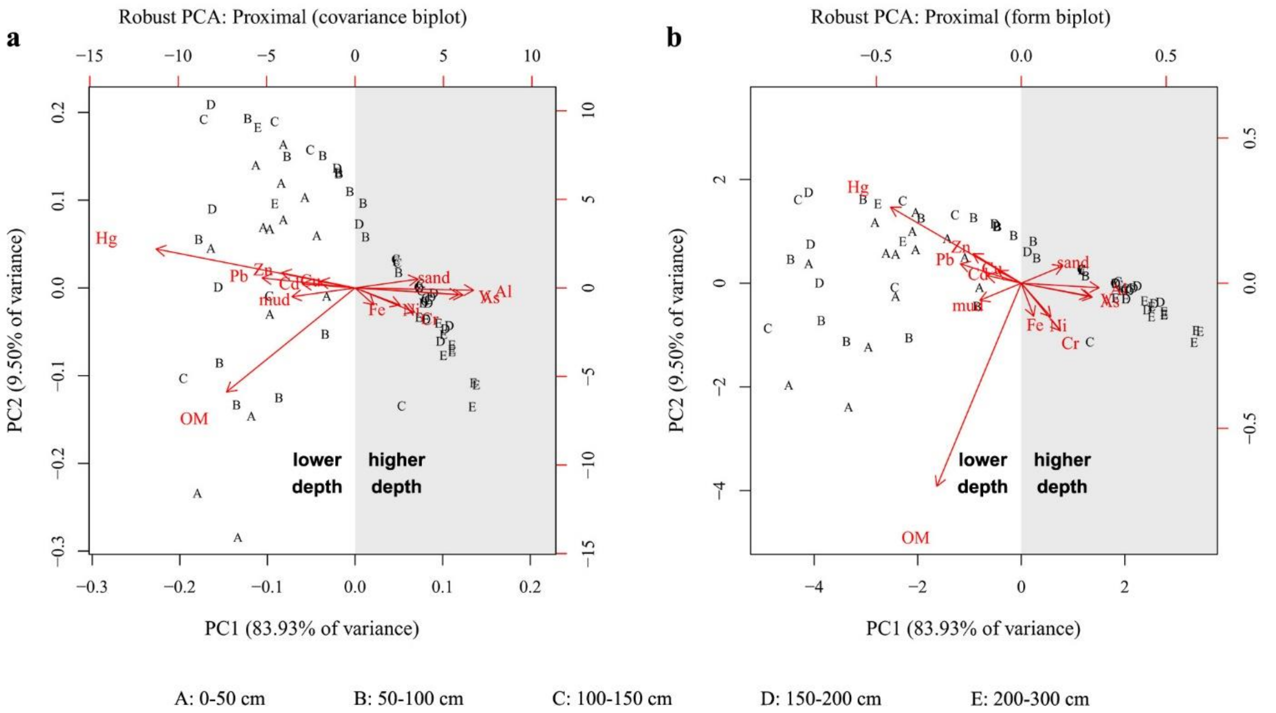


Fig. 17 Robust principal component analysis of the variables analyzed along the selected cores (the bottom right panel of Fig. 16) in the proximal zone: (a) covariance biplot and (b) form biplot. OM is a non-compositional variable whilst the granulometry and elements are two different compositions.

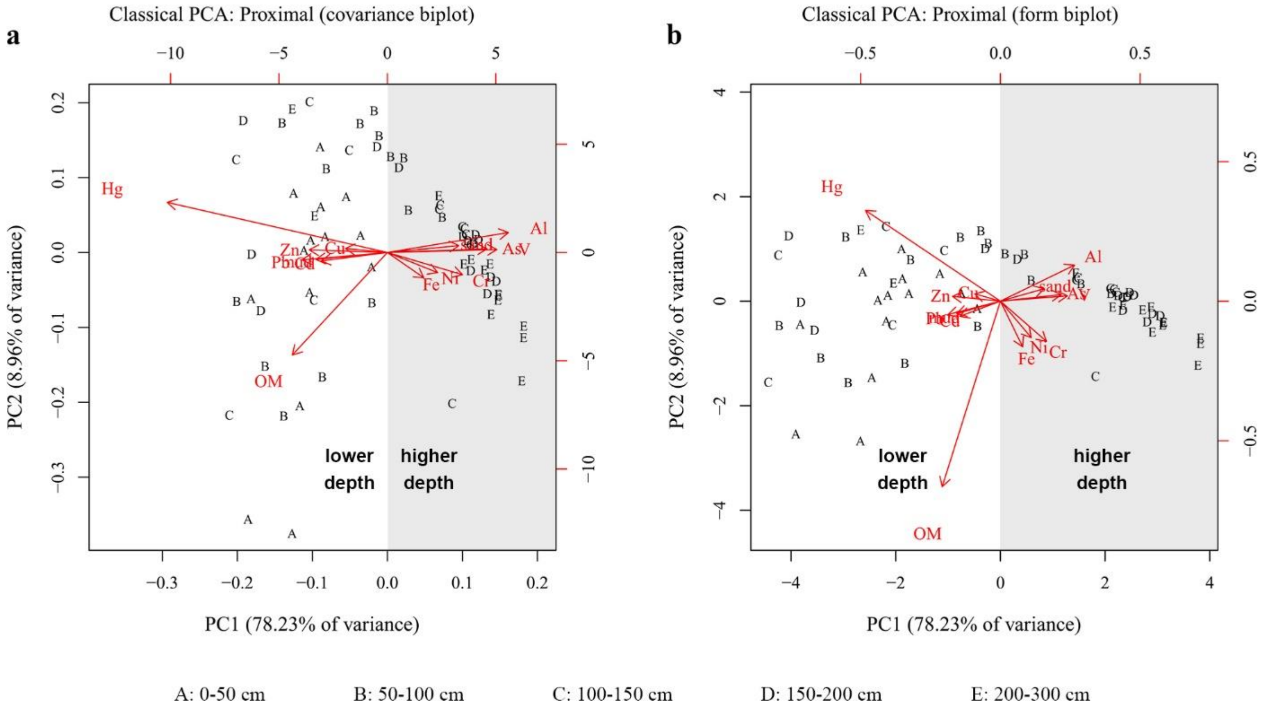


Fig. 18 Classical principal component analysis of the variables analyzed along the selected cores (the bottom right panel of Fig. 16) in the proximal zone: (a) covariance biplot and (b) form biplot. OM is a non-compositional variable whilst the granulometry and element contents are two different compositions.

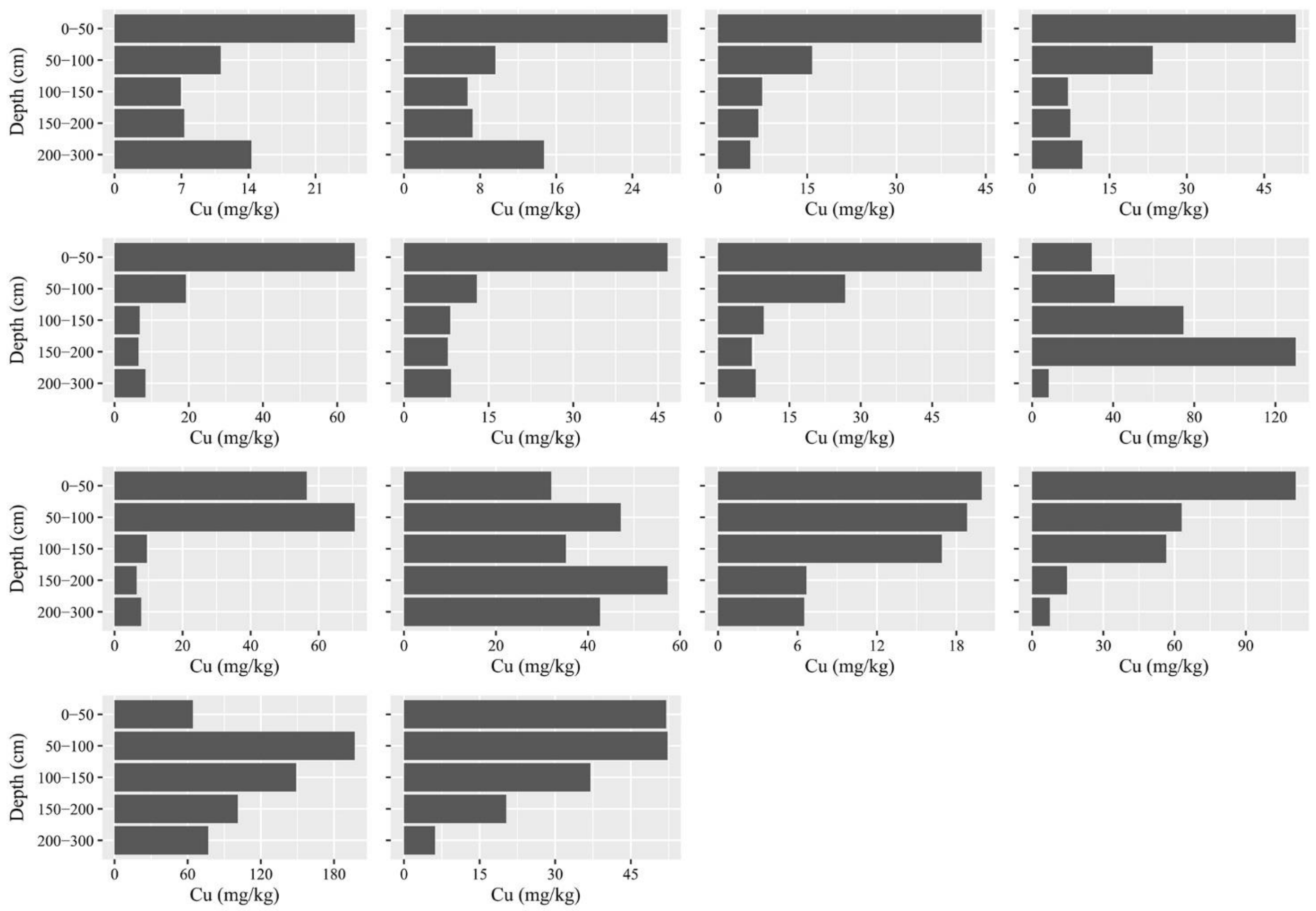


Fig. 19 Copper (mg/kg) profiles of the 14 sediment cores in front of Bagnoli industrial site. Location of the selected cores is represented in the bottom right panel of Fig. 16.

5. Conclusions and future work

The existing literature indicates that geochemistry of the sediments deposited in the eastern sector of Pozzuoli Bay has been affected by natural (e.g., underwater fumaroles, minerals and submarine groundwater discharge) and anthropogenic (for instance, industrial activity in the past, the ineffective coastal landfill and the seepages) factors. Researchers proposed different sources for PTEs (particularly for As) depending on the variables in the dataset and data treatment techniques. In the present investigation, the study area is divided in the proximal and distal zones (PZN and DZN, respectively). In geochemical maps and along the sediment cores, the clr-transformed values of metal(loid)s evidence meaningful univariate patterns. According to the robust PCA of PZN, two strong associations are detected: (1) Hg, Cd, Cu, Pb, Zn, OM, mud, dist_coast and dist_seepage; and (2) Al, As, V, Fe, Cr, Ni, sand and dist_fumarole. The former can be attributed to the industrial activity because concentrations of the PTEs decease along the cores and in the DZN, but the latter is likely contributed by pyroclastic deposits of Phlegrean Fields and the seepages with a hydrothermal component. It seems that submarine groundwater discharge, underwater fumaroles and the coastal landfill do not play the major role in arsenic geochemistry. Despite almost similar associations in DZN, low contribution of variables in the robust PCA probably suggests that the main sources of elements are spatially located in the proximal zone. Therefore, this pioneering work demonstrates the effectiveness of compositional data analysis for bridging the ongoing knowledge gaps in sediment geochemistry of complex coastal environments.

The following future research directions are suggested to shed light on the geochemistry of sea sediments in Pozzuoli Bay: (1) applying the compositional and fractal approaches to the seafloor surface sediments of the same depth; (2) using CoDA technique after collecting spatially dense cores and obtaining high-resolution measurements along them to reduce uncertainty; (3) considering a buffer zone around the Bagnoli-Coroglio site of national interest (SIN), according to the sediment movement path, for the future monitoring plans; (4) investigating chemical forms of PTEs (e.g., the fraction of metals bound in carbonates, Fe-Mn oxides, sulfides, organic matter and lithogenic minerals) in order to define chemical interaction, mobility, bioavailability and potential toxicity of PTEs; and (5) extending the ideas in this chapter to other pollutants such as PAHs.

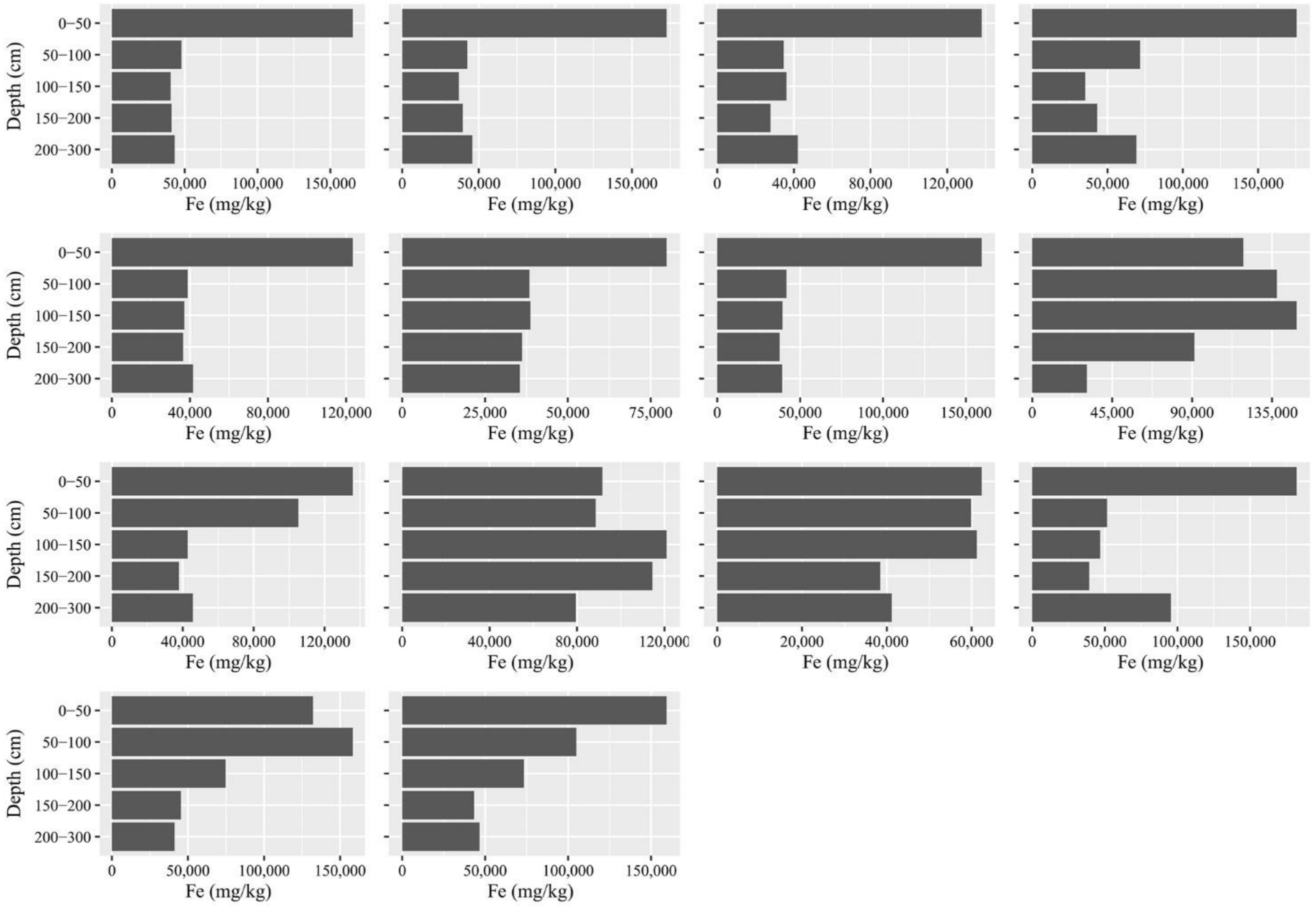


Fig. 20 Iron (mg/kg) profiles of the 14 sediment cores in front of Bagnoli industrial site. Location of the selected cores is represented in the bottom right panel of Fig. 16.

Contribution of the Ph.D. candidate

Determining research objectives; Data visualization; Data interpretation; Writing the original draft; Revising the draft

CHAPTER 5

Soil Geochemistry in Phlegrean Fields

This chapter is in preparation for submission to a journal as a Research Article:

Albanese, S., Ebrahimi, P. *, Aruta, A., Cicchella, D., De Vivo, B., & Lima, A. Potentially toxic elements in the soils of Phlegrean Fields (south Italy) and the immediate surroundings: Spatial distribution, origin and probabilistic human health risk. – *In Preparation*

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

Soil Geochemistry in Phlegrean Fields

Abstract

Phlegrean Fields are an active volcanic area in south Italy where several cities such as Naples, Pozzuoli and Cuma are located. The spatial distribution, source and human health risk of the potentially toxic elements (PTEs) are of growing concern in this area because the population density is high, concentrations of PTEs in soil increased over time and the intensive anthropogenic and volcanic activities might pose adverse health effects to the locals, particularly during the COVID-19 pandemic. To enlighten these points, 394 topsoil (0 to 20 cm) samples are collected and concentrations of PTEs along with some other elements that could signal the origin of PTEs are determined in the <2 mm fraction. Due to the constant sum constraint of geochemical data, the spatial distribution of PTEs is visualized by integrating compositional data analysis and empirical Bayesian kriging which indicates that higher concentrations of Pb, Zn, Cd, Cr, Hg, Ni and Sb are related to the urban area in the east or southeast, but greater levels of As, Tl, Co, Cu, Se and V are determined in the soils from the other parts. For a statistically meaningful dimension reduction, robust principal component analysis is also performed after log-ratio transformation which detected: (1) the Pb–Zn–Hg–Cd–Sb–Cr–Ni association that probably highlights anthropogenic activities such as heavy traffic load and fossil fuel combustion in the urbanized area; (2) the Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association that mostly reveals the contribution of pyroclastic deposits from Phlegrean Fields, Somma–Vesuvius and Roccamonfina volcanic eruptions; and (3) the Na–K–B association that feasibly indicates the weathering degree, being lower in the Neapolitan Yellow Tuff caldera. The probabilistic health risk modelling for the children under 6 years old (i.e. the most sensitive group of community) shows negligible non-carcinogenic risk for exposure to Pb and Zn through soil ingestion (certainty > 95%). Pb carcinogenic risk through the ingestion route is unexpected for all age groups (CR ≤ 1E-06 with above 99% certainty). However, for the inhalation pathway, there is a 90% chance of acceptable (i.e. between 1E-6 and 1E-4) Pb carcinogenic health risk for the children aged from 0 to 1 year old and a 25-45% chance of acceptable cancer risk for those between 1 and 6 years old. This should not be overlooked because Naples is a heavily populated metropolis under high environmental pressure and previous studies reported the increased Pb and Zn quantities in soil over a 26-year timespan. Overall, results of geostatistical interpolation, compositional data analysis and probabilistic health risk modelling are informative for uncovering the link between soil geochemistry and human health.

1. Introduction

Nowadays, the majority of world's population lives in urban environments and the city-based green spaces (such as parks, botanical gardens and greenways) play an important role in outdoor activities and social relationships, especially during a pandemic like the outbreak of COVID-19. Soil composition has attracted public attention worldwide because an increasing number of people are getting aware of the link between human health and the living environment. Conducting enough environmental measurements to reliably address an ongoing scientific problem can be costly and laborious. Further, sampling is not always possible in the desired sample locations. Predicting values where observations are not available is, therefore, highly important and the geostatistical models generally have a better performance than the deterministic spatial interpolation models. Developing a reliable automatic statistical interpolation model has been the main focus of researchers. This is a daunting task because developing each statistical model is dependent on the users' data, and the data complexity cannot be described mathematically without interaction (Krivoruchko, 2012). Recently, Krivoruchko (2012) and Krivoruchko and Gribov (2014, 2019) improved the classical methods and introduced empirical Bayesian kriging (EBK) as a reliable automatic interpolator that creates and then uses many semivariogram models to capture the error in the underlying variogram estimation.

For (geo)statistical analysis, respecting the data structure is also important for reliable extraction and interpretation of information. The measurement units (such as percent, part per million and part per billion) that express the chemical composition of any environmental compartment are vectors of positive values and carry only relative information. It highlights that almost all geochemical data sum up to a constant and have a unique geometry. This type of multivariate data is known as compositional data that needs a transformation to deal with the closure problem for representative statistical analysis in the real space (Aitchison and Greenacre, 2002; Filzmoser et al., 2018; Sadeghi, 2020; Ebrahimi et al., 2022). Aitchison (1982) and Egozcue et al. (2003) proposed additive log-ratio (alr), centered log-ratio (clr) and isometric log-ratio (ilr) transformations in the last decades to equip geochemists with new tools that help find more reliable solutions for the ongoing scientific challenges. Many researchers have recently applied this approach for investigating the composition of rock, soil, sediment and water, and for underlining the geochemical reactions that regulate the main processes in each compartment (e.g., Dmitrijeva et al., 2019; Graziano et al., 2020; Somma et al., 2021; Zissimos et al., 2021). This line of research is still in progress and deserves further attention to interpret a compositional dataset more representatively.

Providing some information on concentrations and sources of PTEs leads to public concern and brings up the potential health issue, in particular for the vulnerable groups. Every individual could be exposed to the pollutants in soil through dermal contact, ingestion and inhalation pathways. Although the micro- and macro-nutrients are essential for the optimum biochemical function of human body, an abnormal condition might occur following excessive exposure to some metals/metalloids (Siegel, 2002; Yang et al., 2019; Huang et al., 2021), depending on their bioaccessibility and bioaccumulation potential. Carcinogenic and non-carcinogenic human health risk could be modelled via deterministic and probabilistic methods. A deterministic approach uses single input values for point estimation of health risk based on a given formula. On the other hand, a probabilistic technique considers the distributions of variables to select a value for each input variable, estimates the health risk based on a given equation, repeats this process for a number of pre-defined runs and returns distribution of the characterized health risk. The latter offers more information to authorities and facilitates risk communication by taking into account the variability and uncertainty in risk, but it needs more time, resources and expertise for application and interpretation (USEPA, 2001).

In Italy, the Domizio-Flegreo Littoral and Agro Aversano (including the study area) have been officially declared as the contaminated sites of national interest (SIN) where industries, intensive agricultural practices and numerous buffalo farms account for the major anthropogenic sources of pollution (Grezzi et al., 2011; Bove et al., 2011). Several researchers (e.g., De Vivo et al., 2016; Petrik et al., 2018a; Thiombane et al., 2019; Guagliardi et al., 2020; Zuzolo et al., 2020) studied soil geochemistry for characterizing origin of PTEs in Campania Region, but human health risk due to exposure to the environmental pollutants has not been extensively assessed with the models proposed by the USEPA (2011). It is worth mentioning that Naples is one of the most populated metropolises in Italy which partly occupies the Phlegrean Fields caldera. Therefore, authorities need to minimize the negative impacts of the urban environment and improve the quality of life which deserves collaboration with scientists to identify the natural vs. anthropogenic contribution. The objectives of the present chapter are to: (1) visualize the spatial distribution of PTEs; (2) characterize the origin of PTEs; and (3) determine the health risk for children. As explained in the abovementioned paragraphs, the recent advances in spatial analysis, geochemical data analysis and health risk assessment will be employed to achieve a comprehensive conclusion for decision makers.

2. Study area

The Campanian Volcanic Zone (CVZ), also known as the Campanian Plain, has an eruptive history of over 600 ka and covers an area of about 2000 km^2 in south Italy (Fig. 1; Rolandi et al., 2003). It is bounded by the Apennines on the east, the Tyrrhenian Sea on the west, the Mt. Massico region on the north, and the Gulf of Naples up to the Island of Capri on the south. Structurally, the CVZ is located within a graben formed in the Mesozoic carbonates of the southern Apennine Mountains that underwent maximum subsidence of 5 km during the Pliocene-Pleistocene (Ippolito et al., 1973). Rolandi et al. (2003) subdivided the CVZ into (Fig. 1): (1) the northern zone, from Mt. Roccamonfina and Mt. Massico to the Lago Patria-Acerra depression, where younger deposits cover the ignimbrite sequence (Rosi and Sbrana, 1987; Romano et al., 1994; Rolandi et al., 2003); (2) the intermediate zone, from west Phlegrean Fields to east Naples, where the ignimbrite sequence together with the underlying sequences of lava and pyroclastic units are exposed as coherent stratigraphic successions (Rosi and Sbrana, 1987; Rolandi et al., 2003); and (3) the southern zone, covering the Bay of Naples up to the Island of Capri and the Sorrentina Peninsula.

Phlegrean Fields is an active volcanic field in the intermediate zone of CVZ that was partially submerged in Pozzuoli Bay. The eruptive events go back to ~80 ka in the study area (Scarpati et al., 2013), but the 40-ka Campanian Ignimbrite (CI) and the 15-ka Neapolitan Yellow Tuff (NYT) are of the main eruptions (De Vivo et al., 2001; Rolandi et al., 2003; Deino et al., 2004; Natale et al., 2021). Although the majority of researchers associated the 12-km-wide Phlegrean Fields caldera with the CI and NYT Pleistocene eruptions and the related caldera collapses (e.g., Rosi and Sbrana, 1987; Fisher et al., 1993; Orsi et al., 1996; Di Vito et al., 1999; Perrotta et al., 2006), some of them such as Scandone et al. (1991), De Vivo et al. (2001), Rolandi et al. (2003) and De Natale et al. (2014, 2016) excluded CI as the caldera-forming eruption. Since the NYT eruption, the centre of Phlegrean Fields caldera spatiotemporally experienced a series of small-volume eruptions with an evolution in the chemistry of erupted products which may suggest the evolutionary trend of a large cooling magma chamber (Armienti et al., 1983). At least 73 phreatomagmatic eruptions younger than NYT occurred within the caldera (Di Vito et al., 1999; Isaia et al., 2012; Scarpati et al., 2013). The eruption of Monte Nuovo volcano was the last one in 1538 CE (Di Vito et al., 1987).

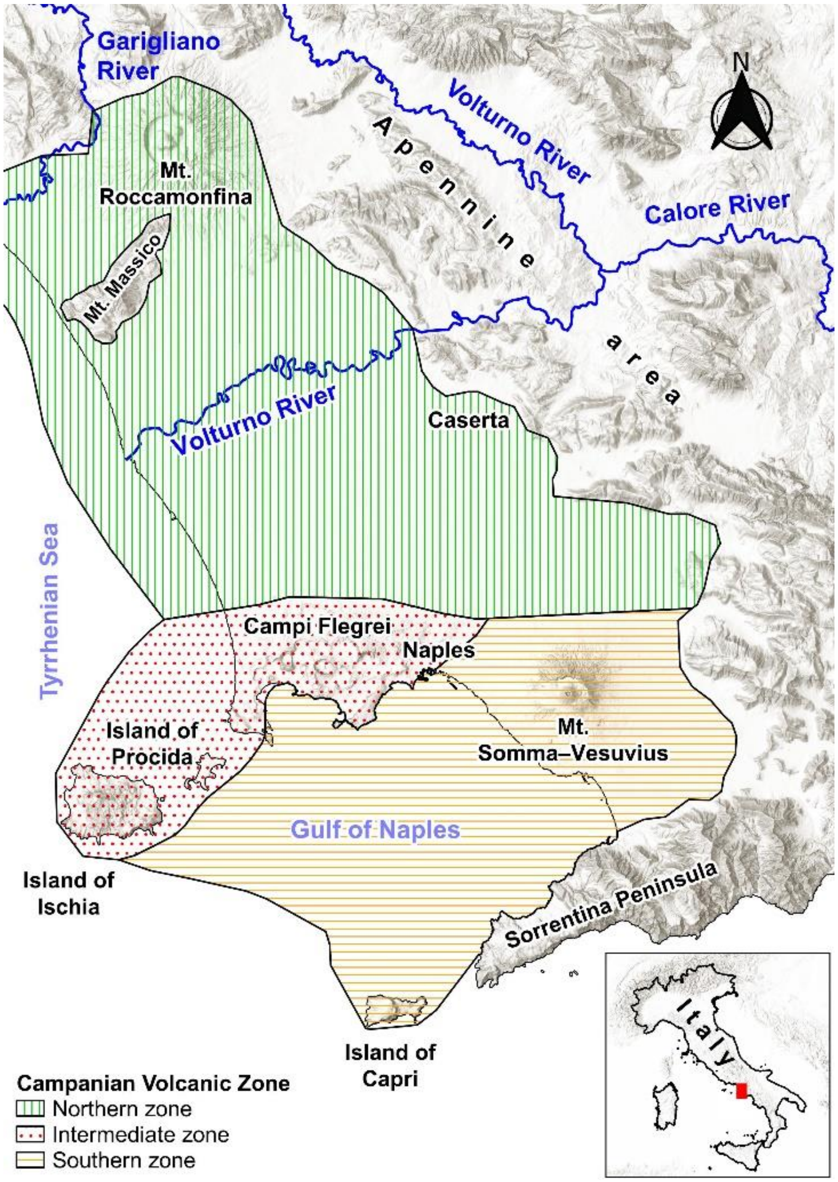


Fig. 1 The Campanian Volcanic Zone and location of the Phlegrean Fields in the intermediate zone (after Rolandi et al., 2003).

In Naples province, the parent material of soil is mostly a combination of Mt. Somma–Vesuvius, Phlegrean Fields and Campanian Ignimbrite volcanic products (Cicchella et al., 2005). Imperato et al. (2003) reported biotite, sanidine, leucite, garnet and pyroxene as the soil mineral assemblage in Naples city and attributed the absence of well-crystallized clay minerals (except for the minor halloysite) to weak pedogenetic processes probably due to continuous provision of fresh volcanic products. Particularly, in Phlegrean Fields, the moderately coarse-textured soil lies on loose pyroclastic deposits (di Gennaro et al., 2002) which is predominantly formed from the trachytic and phonolitic lavas in specific areas, lithoid tuffs (Campanian Ignimbrite and Neapolitan Yellow Tuff) and pyroclastic materials (De Vivo et al., 2006).

3. Materials and methods

3.1. Sampling and analysis

In this study, 394 topsoils (0 to 20 cm) were collected to cover the area inside Phlegrean Fields and its immediate surroundings. The sampling procedure and sample preparation were in accordance with the international guidelines established in the geochemical mapping of agricultural and grazing land soil (GEMAS) project. Briefly, to obtain a representative sample from 0 to 20 cm below the ground surface, the vegetation cover was removed and 1.5 kg of a composite material was collected from vertices and centre of a square (1m × 1m). The supplementary site-specific information such as coordinates, topography, local geology, land use and anthropogenic activity was also included in a filed observation sheet.

In the laboratory, the samples were dried at <35 °C with infrared lamps to avoid mercury volatilization before passing them from a 2-mm sieve. The fraction finer than 2 mm contains the majority of reactive surfaces in soil. After aqua regia digestion, the "pseudo-total" concentrations of 53 elements were determined using a combination of inductively coupled plasma-atomic emission spectrometry (ICP-AES) and inductively coupled plasma-mass spectrometry (ICP-MS). The incomplete dissolution of soil mineral matrix using aqua regia results in the measurement of pseudo-total content of elements which is useful for evaluating the potential risk (Gupta et al., 1996). The relative percentage difference was computed to check quality of the analytical results. Precision ranges from ± 100% at the detection limit to better than ± 10% at concentrations ≥50 times the detection limit. In this study, only potentially toxic elements (As, Cd, Co, Cr, Cu, Hg, Ni, Pb, Sb, Se, Tl, V and Zn) and other elements such as Al, B, Fe, K, Mn, Na, Th, Ti and U that probably provide some information on the origin of PTEs are of interest.

3.2. Geochemical mapping

The empirical Bayesian kriging (EBK) is used for spatial analysis in this chapter. It is an automatic interpolator that captures the error in the underlying variogram estimation as follows (Krivoruchko, 2012; Krivoruchko and Gribov, 2014, 2019): (1) estimating a semivariogram model from the data; (2) simulating a new value at each input data location using this semivariogram; (3) estimating a new semivariogram model from the simulated data and calculating a weight for the corresponding semivariogram using Bayes' rule, which reveals likeliness of generating the observed data from the semivariogram; and (4) repeating steps 2 and 3 to simulate a new set of values at the input locations, estimate a new semivariogram model and its weight for each repetition. The semivariogram distribution and the obtained weights are finally used for producing predictions and prediction standard errors at the unsampled locations. Compared with the classical kriging models, EBK is more reliable while the disadvantages are: (1) limitations for large datasets due to estimating the variogram parameters via restricted maximum likelihood (Basu, 2016); (2) slower performance, particularly for exporting the results in raster; and (3) unavailability of cokriging and anisotropy (ESRI, 2016).

In the cross-validation process, the predicted and the observed values are also compared via Mean Error (ME), Root Mean Square Error (RMSE), Average Standard Error (ASE), Mean Standardized Error (MSE) and Root Mean Square Standardized Error (RMSSE):

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{2}$$

$$\mathrm{ASE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sigma^2(y_i)} \tag{3}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\frac{(y_i - x_i)}{\sigma(y_i)} \tag{4}$$

$$\mathrm{RMSSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\frac{(y_i - x_i)}{\sigma(y_i)}\right]^2} \tag{5}$$

where y_i, x_i and σ represent the simulated value, the observed value and the standard deviation, respectively. Cross-validation is the process of extracting an individual data point and predicting its value based on the remaining information at the rest of locations. Prior to continuing with the next point, the extracted observation is restored. After prediction for all data points, some rules are applied to check if the variability of predictions is correctly assessed, overestimated or underestimated (Kresic and Mikszewski, 2012; Li et al., 2020).

In this study, Esri ArcMap10.8 is utilized for generating the EBK models with raw and clr-transformed data to check whether addressing data closure improves the predictions. The clr transformation refers to dividing a given compositional part by the geometric mean of all parts in the same observation, and taking the natural logarithm (Aitchison, 1982). The K-Bessel semivariogram, the empirical transformation, 500 iterations, the overlap factor of 1 and the subset size of 100 are considered for the analysis. The Search Neighborhood parameter is also set to standard circular with a search radius of 1000 m (min and max neighbors are 5 and 10, respectively). The spatial maps are finally classified based on the concentration-area (C-A) fractal analysis. For more information on the classification procedure, the interested readers are referred to Cheng et al. (1994) and Somma et al. (2021).

3.3. Multivariate analysis

3.3.1. Correlation diagram

Krumbein and Graybill (1965) proposed the correlation diagram as a graphical tool for representing the associations between variables. Sinclair and Blackwell (2002) confirmed that applying this tool for grouping the correlated variables is informative. A correlation diagram is constructed in the following steps (de Campos and Licht, 2021): (1) generating a correlation matrix; (2) finding the pair of variables with the highest significant positive correlation coefficient; (3) assigning a circle to each variable and putting the corresponding correlation coefficient over the line connecting the circles; (4) choosing another pair of variables with a relatively lower correlation coefficient for drawing the pairs and their connecting line; (5) connecting the latest set of variables with the previously drawn pairs, if a significant correlation exists; and (6) repeating steps 4 and 5 until reaching a pre-defined cutoff value for the correlation coefficients, determined according to the background knowledge of user. Because of the constant sum problem in geochemical data, symmetric coordinates of each pair of variables are computed (further information is available in Filzmoser et al. (2018) and Somma et al. (2021)), the spearman correlation matrix (Table 1) is obtained and a threshold of 0.4 is considered for generating the correlation diagram in this chapter. The EzCorrGraph app, developed by de Campos and Licht (2021), is used for the visualization of correlation matrix.

Table 1 Spearman correlation coefficients between the variables in this study. To deal with data closure problem, the symmetric coordinates were calculated prior to computing the correlation coefficients. The values ≤ |±0.4| are underlined.

Element	**Cu**	**Pb**	**Zn**	**Ni**	**Co**	**Mn**	**Fe**	**As**	**U**	**Th**	**Cd**
Cu	1.00	0.24	0.29	-0.22	-0.31	-0.25	-0.27	-0.06	0.05	-0.20	0.11
Pb	0.24	1.00	0.60	0.40	-0.37	-0.42	-0.48	-0.33	-0.51	-0.48	0.41
Zn	0.29	0.60	1.00	0.28	-0.43	-0.45	-0.43	-0.44	-0.50	-0.52	0.57
Ni	-0.22	0.40	0.28	1.00	0.31	-0.18	-0.07	-0.21	-0.56	-0.37	0.33
Co	-0.31	-0.37	-0.43	0.31	1.00	0.54	0.74	0.33	0.22	0.35	-0.37
Mn	-0.25	-0.42	-0.45	-0.18	0.54	1.00	0.72	0.64	0.60	0.70	-0.33
Fe	-0.27	-0.48	-0.43	-0.07	0.74	0.72	1.00	0.47	0.42	0.53	-0.36
As	-0.06	-0.33	-0.44	-0.21	0.33	0.64	0.47	1.00	0.61	0.59	-0.34
U	0.05	-0.51	-0.50	-0.56	0.22	0.60	0.42	0.61	1.00	0.80	-0.42
Th	-0.20	-0.48	-0.52	-0.37	0.35	0.70	0.53	0.59	0.80	1.00	-0.46
Cd	0.11	0.41	0.57	0.33	-0.37	-0.33	-0.36	-0.34	-0.42	-0.46	1.00
Sb	0.03	0.72	0.58	0.45	-0.37	-0.50	-0.50	-0.43	-0.67	-0.50	0.53
V	-0.28	-0.50	-0.48	-0.03	0.80	0.56	0.86	0.37	0.35	0.42	-0.47
Cr	0.13	0.37	0.43	0.62	-0.05	-0.32	-0.25	-0.29	-0.44	-0.47	0.42
Ti	-0.18	-0.52	-0.53	-0.19	0.65	0.73	0.79	0.56	0.65	0.77	-0.46
B	-0.20	-0.30	-0.05	-0.12	0.00	-0.03	0.00	0.01	0.13	0.02	-0.07
Al	-0.17	-0.57	-0.53	-0.37	0.53	0.73	0.75	0.48	0.75	0.79	-0.47
Na	-0.10	-0.47	-0.22	-0.49	0.05	0.12	0.21	0.06	0.32	0.28	-0.34
K	-0.13	-0.43	-0.32	-0.46	0.20	0.28	0.31	0.07	0.41	0.38	-0.45
Tl	-0.07	-0.51	-0.54	-0.36	0.44	0.63	0.59	0.58	0.71	0.70	-0.39
Hg	0.07	0.64	0.42	0.41	-0.28	-0.43	-0.49	-0.32	-0.51	-0.49	0.29
Se	0.09	-0.09	-0.04	-0.10	-0.16	-0.04	-0.20	0.06	0.03	-0.02	0.15

Table 1 (continued)

Element	**Sb**	**V**	**Cr**	**Ti**	**B**	**Al**	**Na**	**K**	**Tl**	**Hg**	**Se**
Cu	0.03	-0.28	0.13	-0.18	-0.20	-0.17	-0.10	-0.13	-0.07	0.07	0.09
Pb	0.72	-0.50	0.37	-0.52	-0.30	-0.57	-0.47	-0.43	-0.51	0.64	-0.09
Zn	0.58	-0.48	0.43	-0.53	-0.05	-0.53	-0.22	-0.32	-0.54	0.42	-0.04
Ni	0.45	-0.03	0.62	-0.19	-0.12	-0.37	-0.49	-0.46	-0.36	0.41	-0.10
Co	-0.37	0.80	-0.05	0.65	0.00	0.53	0.05	0.20	0.44	-0.28	-0.16
Mn	-0.50	0.56	-0.32	0.73	-0.03	0.73	0.12	0.28	0.63	-0.43	-0.04
Fe	-0.50	0.86	-0.25	0.79	0.00	0.75	0.21	0.31	0.59	-0.49	-0.20
As	-0.43	0.37	-0.29	0.56	0.01	0.48	0.06	0.07	0.58	-0.32	0.06
U	-0.67	0.35	-0.44	0.65	0.13	0.75	0.32	0.41	0.71	-0.51	0.03
Th	-0.50	0.42	-0.47	0.77	0.02	0.79	0.28	0.38	0.70	-0.49	-0.02
Cd	0.53	-0.47	0.42	-0.46	-0.07	-0.47	-0.34	-0.45	-0.39	0.29	0.15
Sb	1.00	-0.50	0.41	-0.60	-0.12	-0.69	-0.29	-0.38	-0.63	0.60	0.03
V	-0.50	1.00	-0.24	0.73	0.13	0.66	0.30	0.45	0.53	-0.43	-0.27
Cr	0.41	-0.24	1.00	-0.33	-0.17	-0.46	-0.41	-0.48	-0.46	0.29	-0.14
Ti	-0.60	0.73	-0.33	1.00	-0.09	0.87	0.15	0.30	0.77	-0.61	-0.17
B	-0.12	0.13	-0.17	-0.09	1.00	0.04	0.67	0.52	0.04	-0.05	0.14
Al	-0.69	0.66	-0.46	0.87	0.04	1.00	0.29	0.48	0.81	-0.63	-0.05
Na	-0.29	0.30	-0.41	0.15	0.67	0.29	1.00	0.80	0.24	-0.25	0.08
K	-0.38	0.45	-0.48	0.30	0.52	0.48	0.80	1.00	0.39	-0.26	-0.04
Tl	-0.63	0.53	-0.46	0.77	0.04	0.81	0.24	0.39	1.00	-0.58	0.04
Hg	0.60	-0.43	0.29	-0.61	-0.05	-0.63	-0.25	-0.26	-0.58	1.00	0.06
Se	0.03	-0.27	-0.14	-0.17	0.14	-0.05	0.08	-0.04	0.04	0.06	1.00

3.3.2. Robust principal component analysis (PCA)

The prime purpose of principal component analysis (PCA) is to summarize and reduce dimensionality of multivariate data structure. A smaller number of variables (i.e. principal components in this case) are identified in PCA that contain the most important information in the larger dataset in a certain sense. To this aim, the clr transformation was introduced as a preferable transformation (Aitchison and Greenacre, 2002) for compositional data. However, data outliers avoid reliable estimation of the principal components (Maronna et al., 2006) in classical PCA because it is based on the classical sample covariance matrix. Robust PCA is an alternative that benefits from a robust estimation of the covariance matrix (Filzmoser, 1999). This procedure, however, applies the ilr transformation (Egozcue et al., 2003) instead of the clr transformation and back-transforms the resulting scores and loadings to the clr space after robust covariance estimation (Filzmoser et al., 2018). Because it has been indicated (e.g., Filzmoser et al., 2009; Somma et al., 2021) that performance of robust PCA is better than classical PCA, only the former is performed in this study via the R package “robcompositions” (Templ et al., 2011).

3.4. Human health risk modelling

USEPA (2011) proposed a method to evaluate the carcinogenic and non-carcinogenic human health risks associated with the exogenous compounds. Carcinogenic risk refers to the probability of developing cancer over a lifetime due to exposure to either a specific environmental contaminant or a mixture of them (Kamarehie et al., 2019). In contrast, non-carcinogenic risk corresponds to other effects such as mutagenicity, developmental toxicity, neurotoxicity and reproductive toxicity. For non-carcinogenic health risk due to exposure through ingestion route, the average daily doses (ADD, mg/kg day^{-1}) of Pb and Zn are separately calculated as follows:

$$\mathrm{ADD_{ing}} = \frac{\mathrm{C \times IngR \times EF \times ED}}{\mathrm{BW \times AT_{nc}}} \tag{6}$$

and compared with the corresponding reference dose (RfD; mg/kg day^{-1}) to obtain the hazard quotient (HQ):

$$\mathrm{HQ_{ing}} = \frac{\mathrm{ADD_{ing}}}{\mathrm{RfD_{ing}}} \tag{7}$$

where C is the pollutant concentration in soil (mg/kg), IngR is the soil ingestion rate (mg/day), EF is the exposure frequency (day), ED is the exposure duration (year; in this investigation, 1 year for the age groups in section 4.3), BW is the body weight of the exposed individual (kg), $\mathrm{AT_{nc}}$ is the average life span (ED × 365 days). An HQ ≤ 1 means negligible hazard, but HQ > 1 shows the increasing potential for adverse non-carcinogenic effects.

On the other hand, for Pb carcinogenic health risk assessment, ADD of ingestion and inhalation pathways ($\mathrm{ADD_{ing}}$ and $\mathrm{ADD_{inh}}$, respectively) are computed via Eqs. 8 and 9:

$$\mathrm{ADD_{ing}} = \frac{\mathrm{C \times IngR \times EF \times ED}}{\mathrm{BW \times AT_{ca}}} \tag{8}$$

$$\mathrm{ADD_{inh}} = \frac{\mathrm{C \times InhR \times EF \times ED}}{\mathrm{PEF \times BW \times AT_{ca}}} \tag{9}$$

and multiplied by the cancer slope factor ($\mathrm{CSF_i}$) to obtain cancer risk ($\mathrm{CR_i}$) for each exposure route:

$$\mathrm{CR_i} = \mathrm{ADD_i} \times \mathrm{CSF_i} \tag{10}$$

where i refers to the exposure pathway. The parameters in the Eqs. 8 and 9 are the same with Eq. (6), except for the average life span ($\mathrm{AT_{ca}}$ = life expectancy × 365 days), inhalation rate (InhR) and particulate emission factor (PEF = 1.36×10^9 m^3/kg). Human suffers from significant cancer risk if CR ≥ 1E-04, but the values between 1E-6 and 1E-4 refer to the acceptable risk range. Negligible risk to human health is generally expected when CR ≤ 1E-06.

The abovementioned health risk assessment approach traditionally relies on some deterministic parameters that causes either underestimation or overestimation of the risk level and also restricts accurate identification of the riskiest PTE due to ignoring the underlying uncertainty (Yang et al., 2019). The Monte Carlo simulation has proved to be one of the most useful techniques for dealing with the shortcomings (Tong et al., 2018; Aruta et al., 2020) and it is therefore implemented in this chapter using the Oracle Crystal Ball software (version 11.1, Oracle, Inc., USA) for more representative human health risk assessment.

4. Results and discussion

4.1. Soil composition

A summary of the chemical data obtained from instrumental analysis of the collected soil samples is presented in Table 2. The median values of PTEs show that concentrations of Zn, Cu, Pb, V and As are greater (>10 mg/kg) than Cr, Co, Ni, Tl, Sb, Se, Cd and Hg. Some other elements (i.e. Al, Fe, Mn, Na, K, B, Ti, U and Th) are also included in Table 2 which might help delineate source of the PTEs. Of these elements, aluminum, iron and potassium contents are at least one order of magnitude greater (median > 10000 mg/kg). The median values of Mn, Na and Ti range between 500 to 4000 mg/kg, but those of B, U and Th are below 15 mg/kg.

The geochemical maps of PTEs generated by EBK are shown in Fig. 2. Although the spatial patterns did not significantly change (except for Co and V) after dealing with data closure, the cross-validation results reveal that visualizing the spatial distribution of PTEs after clr-transformation is more reliable in most cases (Table 3). It is worth mentioning that all variables listed in Table 2 are considered for log-ratio transformation of data prior to spatial analysis. Visual inspection of the maps indicates that higher concentrations of Pb, Zn, Cd, Cr, Hg, Ni and Sb occur in the eastern or southeastern sector of the study area, but greater levels of As, Tl, Co, Cu, Se and V are determined in the soils from west (Fig. 2).

Table 2 Descriptive statistics for chemical composition of the soil samples (n = 394) in Phlegrean Fields and the immediate surroundings. Units are in mg/kg, unless otherwise specified. P5, P25, P75 and P95 denote the 5th, 25th, 75th and 95th percentiles, respectively.

Variable	**Element**	**D.L.**	**Min**	**P5**	**P25**	**Median**	**Average**	**P75**	**P95**	**Max**	**S.D.**	**MAD**	**IQR**	**CVR (%)**
Potentially toxic elements (PTEs)	**As**	0.1	3.6	7.9	10.2	12.1	14.2	15.4	23.0	163.8	11.4	2.4	5.2	19.8
	Cd	0.01	<0.01	0.09	0.16	0.21	0.32	0.32	0.87	6.18	0.44	0.07	0.16	33.3
	Co	0.1	2.0	3.0	4.4	5.1	5.5	6.0	8.3	36.6	3.1	0.8	1.6	15.7
	Cr	0.5	0.8	2.3	3.8	5.8	8.6	9.2	21.2	149.1	12.1	2.2	5.4	37.9
	Cu	0.01	2.5	15.6	36.7	68.4	92.3	118.4	249.8	465.4	81.1	39.3	81.7	57.4
	Hg	0.005	0.01	0.02	0.03	0.06	0.13	0.13	0.47	2.66	0.23	0.04	0.10	57.9
	Ni	0.1	0.5	2.2	3.2	4.4	6.1	6.7	13.0	100.9	7.7	1.5	3.5	34.1
	Pb	0.01	11.3	24.9	42.4	59.0	97.7	92.4	270.6	2052.2	152.2	20.5	49.9	34.8
	Sb	0.02	0.1	0.3	0.5	0.7	1.4	1.2	4.3	40.8	2.7	0.3	0.7	40.3
	Se	0.1	<0.1	0.07	0.10	0.30	0.29	0.40	0.60	1.70	0.20	0.10	0.30	33.3
	Tl	0.02	0.18	0.80	1.00	1.20	1.24	1.42	1.94	2.58	0.35	0.22	0.42	18.3
	V	2	20.0	29.0	39.0	45.0	45.8	50.8	65.0	187.0	13.4	6.0	11.8	13.3
	Zn	0.1	25.1	45.5	66.0	90.7	122.9	132.4	297.4	1282.8	124.0	27.4	66.4	30.2
Other elements	**Al**	100	9800.0	20295.0	27425.0	31950.0	33165.7	38825.0	47770.0	65700.0	8409.0	5450.0	11400.0	17.1
	B	1	4.0	6.0	8.0	10.0	11.8	13.0	22.0	52.0	6.2	2.0	5.0	20.0
	Fe	100	8600.0	12300.0	16000.0	18500.0	18988.3	20600.0	24435.0	154600.0	8503.9	2250.0	4600.0	12.2
	K	100	2800.0	7195.0	9525.0	11300.0	11890.9	13500.0	18080.0	39900.0	3938.4	1900.0	3975.0	16.8
	Mn	1	240.0	380.6	514.0	595.0	644.0	699.5	897.2	5923.0	383.5	89.0	185.5	15.0
	Na	10	440.0	2093.0	3062.5	4000.0	4461.2	5360.0	8409.5	18300.0	2183.0	1110.0	2297.5	27.8
	Th	0.1	3.3	5.9	8.5	11.0	11.6	14.2	18.6	30.0	4.2	2.9	5.7	26.0
	Ti	10	490.0	670.0	860.0	1020.0	1046.5	1195.0	1530.0	2090.0	262.8	160.0	335.0	15.7
	U	0.1	0.8	2.1	2.6	3.2	3.5	4.0	6.1	10.1	1.3	0.7	1.4	21.9

D.L.: Detection limit; S.D.: Standard deviation; MAD: Median absolute deviation; IQR: Interquartile range; CVR: Robust coefficient of variation

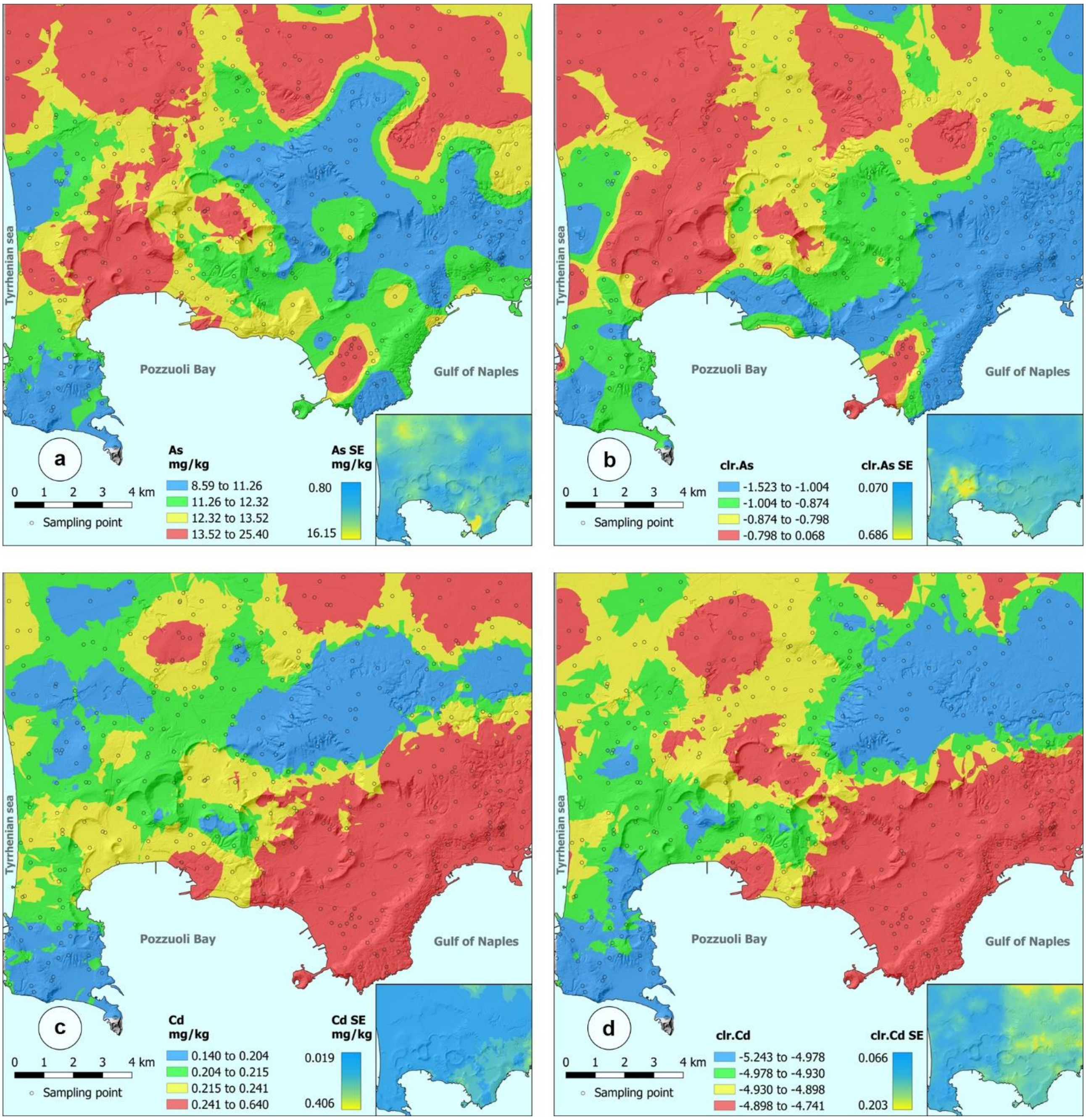


Fig. 2 Spatial distribution of As (a and b), Cd (c and d), Co (e and f), Cr (g and h), Cu (i and j), Hg (k and l), Ni (m and n), Pb (o and p), Sb (q and r), Se (s and t), Tl (u and v), V (w and x), and Zn (y and z) in the study area. The left and right geochemical maps are generated by raw and clr-transformed values, respectively. The lower right insets are the prediction standard error maps.

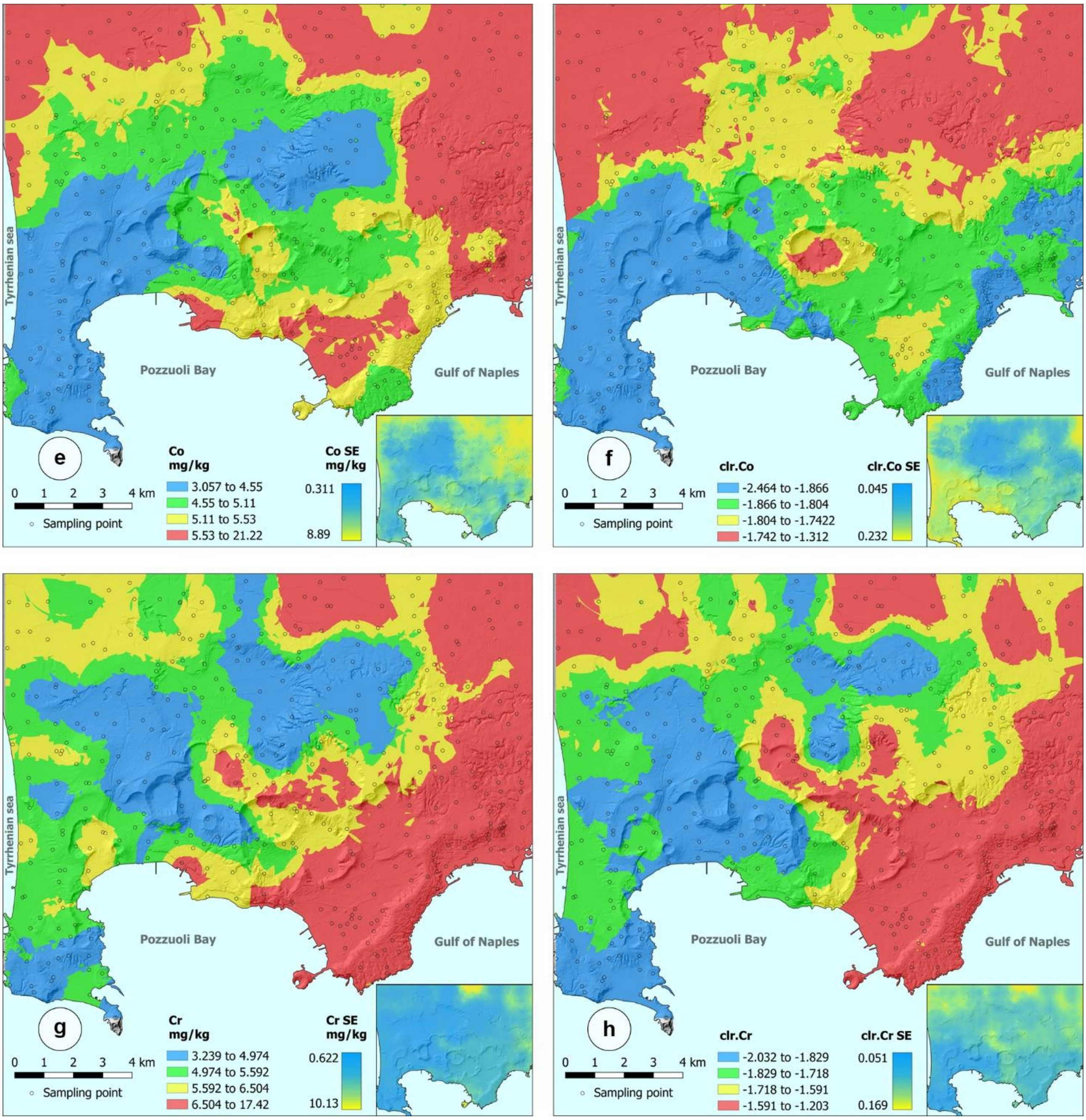


Fig. 2 (continued)

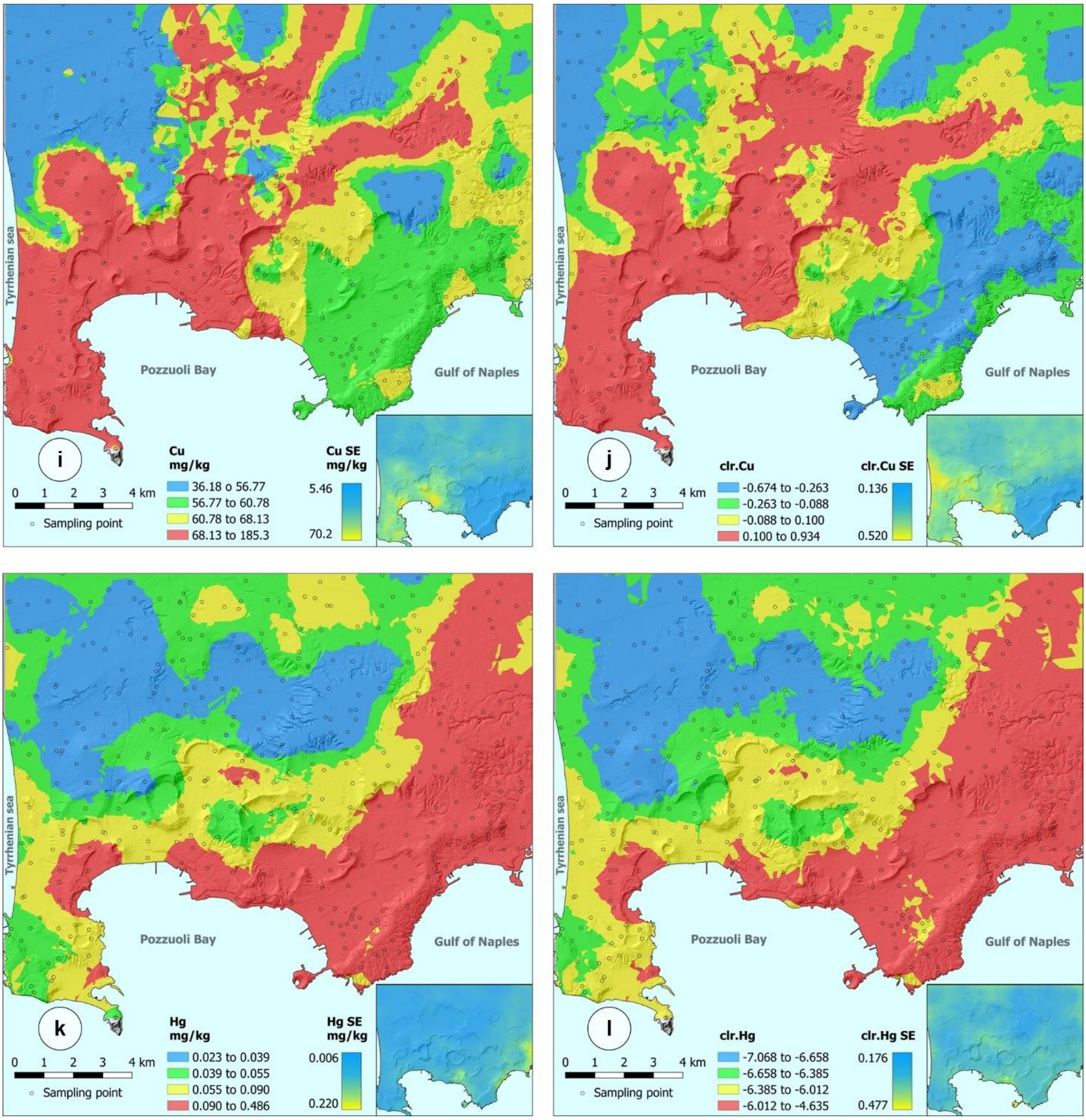


Fig. 2 (continued)

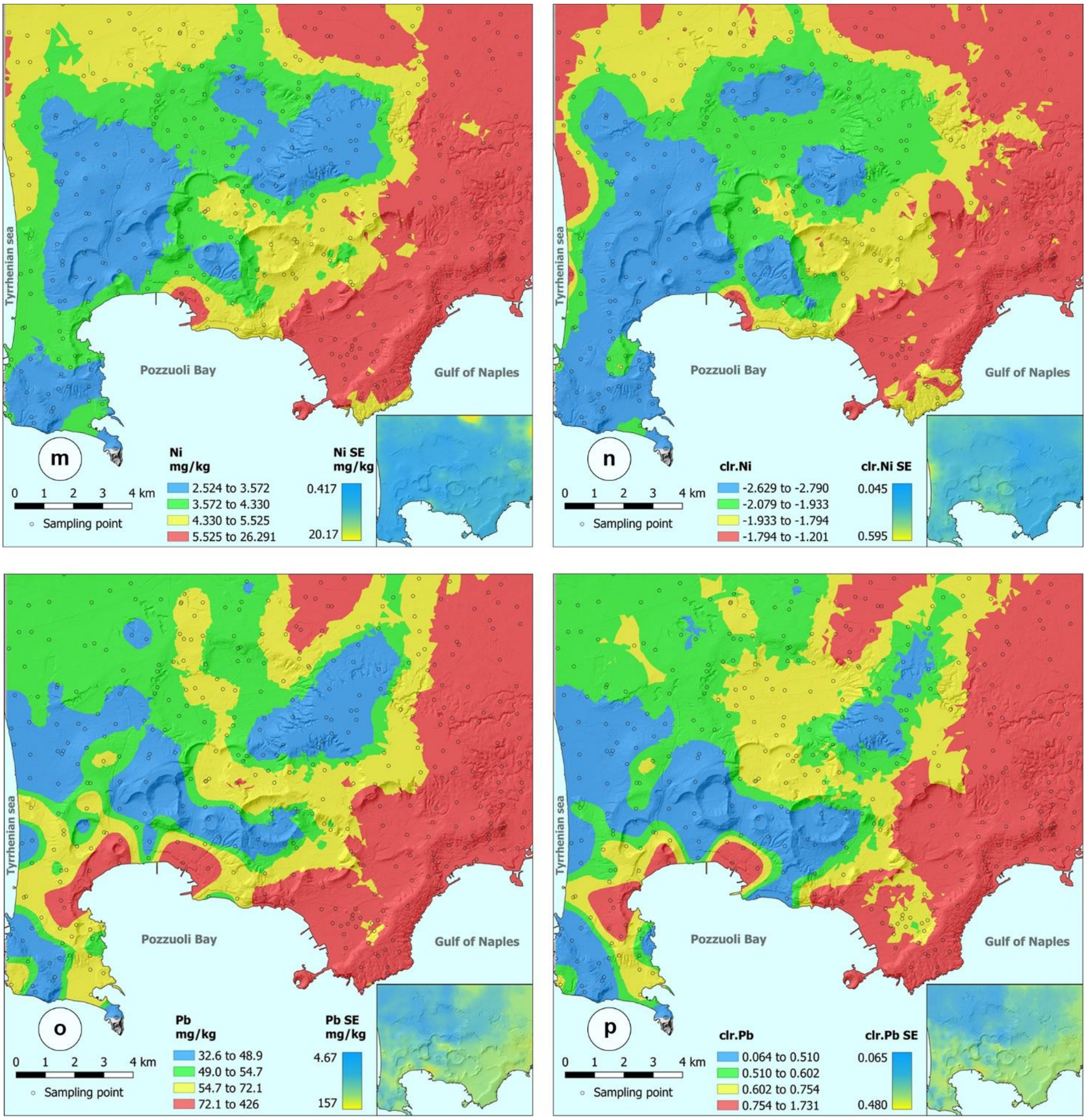


Fig. 2 (continued)

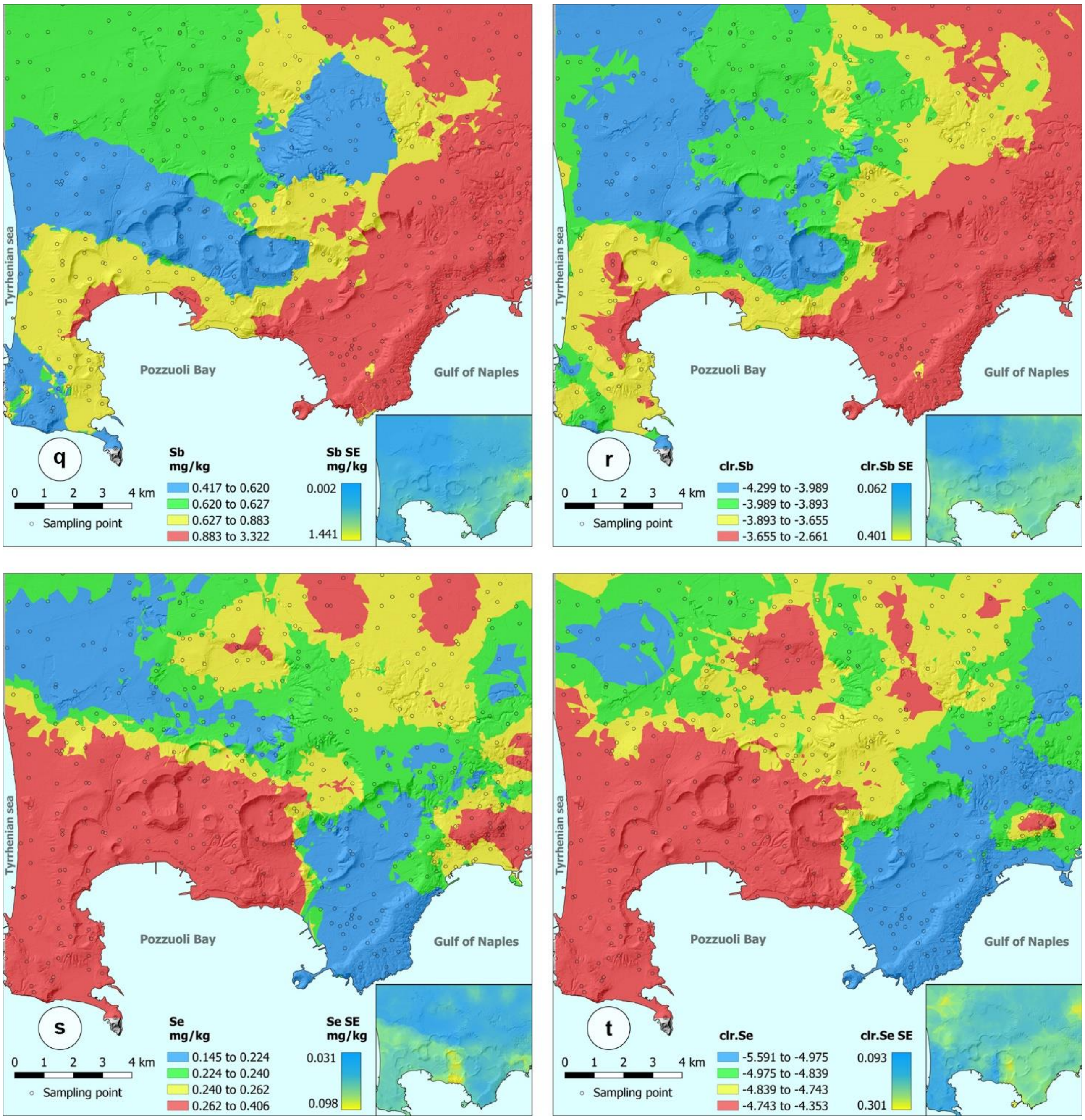


Fig. 2 (continued)

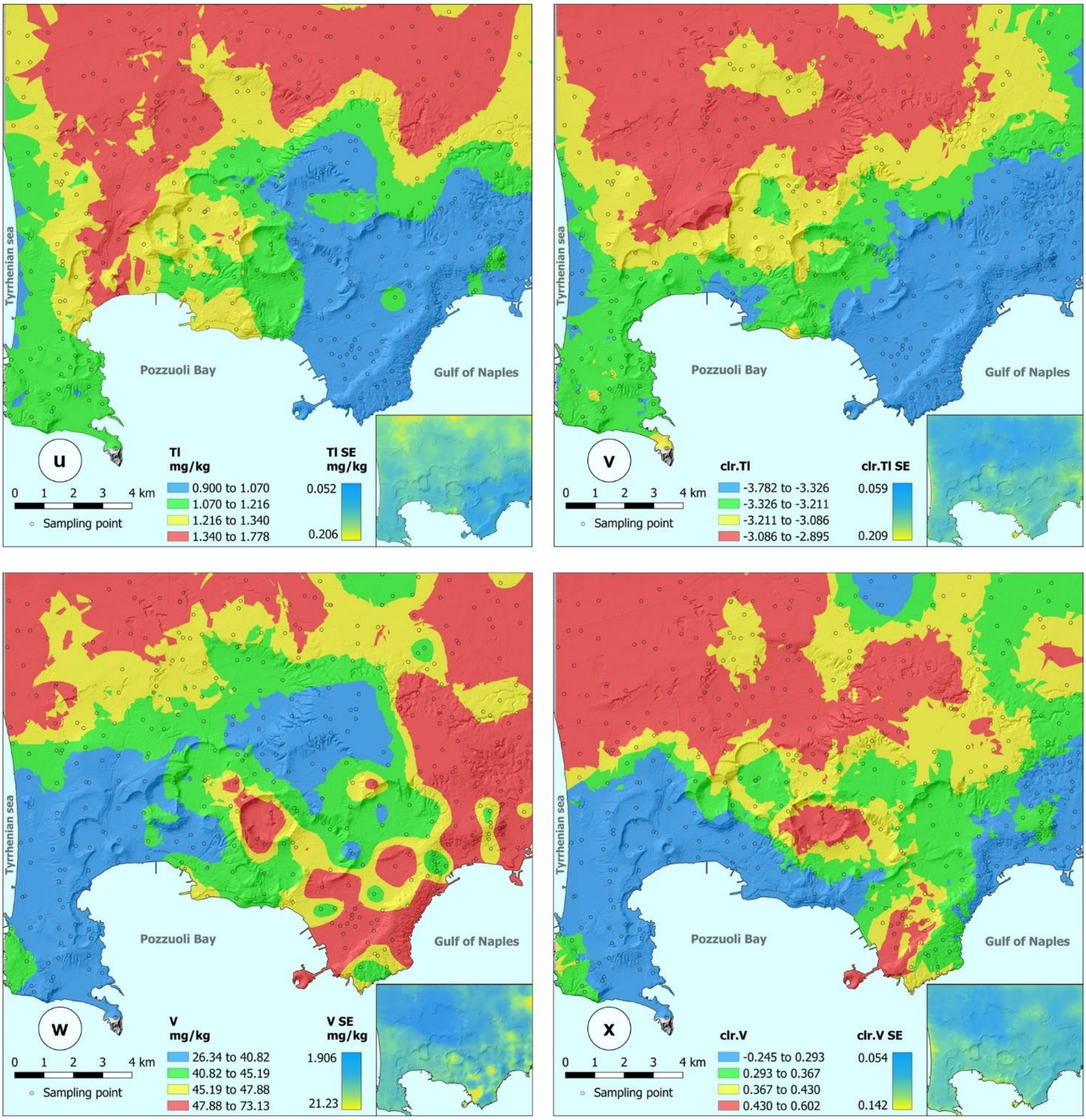


Fig. 2 (continued)

4.2. Probable sources of PTEs

Unique groups of PTEs have been regularly identified via statistical techniques to facilitate distinguishing their source based on the common geochemical behaviour of the members in each group (Keshavarzi et al., 2015; Zhang et al., 2020). The correlation diagram of the elements (Fig. 3) demonstrates three associations: A1) Pb, Zn, Hg, Cd, Sb, Cr and Ni; A2) Al, Fe, Mn, Ti, Tl, V, Co, As, U and Th; and A3) Na, K and B. Copper and selenium which are uncorrelated with other elements (Table 1) are omitted for source identification, but Thiombane et al. (2019) linked the Cu–P relationship with the application of pesticides and phosphate fertilizers in agriculture practices. The spatial distribution of copper is generally irrelevant to the urban area of Naples (Fig. 2i and j). Robust PCA is also performed to confirm the geochemical associations. The first two principal components (PCs) explain a total variance of ca. 70% and their biplot will be informative for analyzing the multivariate dataset (Fig. 4). The members of first two groups (i.e. A1 and A2 in Fig. 3) are correlated with PC1 which captures roughly 59% of the variation in dataset although sodium, potassium and boron correspond to PC2 with about 12% of the variance.

The Pb–Zn–Hg–Cd–Sb–Cr–Ni association probably highlights the anthropogenic origin. Length of the Hg ray is the longest (Fig. 4) which expresses the dominance of this compositional variable. An investigation on the elements with potential human health risk revealed elevated baseline values for Pb, Zn, Sb, Hg, Cd and Cr in the volcanic soils of Naples province and confirmed the evident anthropogenic (such as the heavy traffic load and fossil fuel combustion) contribution (Imperato et al., 2003; Cicchella et al., 2005). The Hg concentration could be partially related to hydrothermal activity as well (Cicchella et al., 2005; De Vivo et al., 2016). In the soils from Giugliano (north Phlegrean Fields), the ^{208}Pb$^{/206}$Pb and ^{207}Pb$^{/206}$Pb values up to 50 cm depth share similarities with those of the Bagnoli industrial site, gasoline, AZIN 30 PB pesticide and anthropogenic aerosols, but they approach the natural values in Mt. Somma–Vesuvius rocks and Neapolitan Yellow Tuff with increasing depth (Grezzi et al., 2011). Regarding the high sampling density in this chapter, the hotspot area for lead is located in the east and southeast (Fig. 2) and lead contamination in north Phlegrean Fields is most likely contributed by some local point sources.

The Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association likely reveals the natural origin. The shorter length of Fe, Mn, Co and V rays explains the lower variation of these components. The strong association of Al, Ti, Tl, As, V and Th probably refers to the immobile elements during weathering of the pyroclastic deposits from Phlegrean Fields, Roccamonfina and Somma–Vesuvius volcanic eruptions (De Vivo et al., 2010, 2016; Thiombane et al., 2019). Therefore, it is expected that Al, Fe, Mn, Ti, Tl, V, Co, As, U and Th mostly come from the geogenic source and occur in the residual fraction of soil. Cicchella et al. (2005) made a relationship between the elevated As baseline values (from 30 to 60 mg/kg) in west Phlegrean Fields and geothermal activity in the Mofete fields, where thermal water has been used for bathing since Roman times. The spatial patterns of PTEs also support the results (Fig. 2).

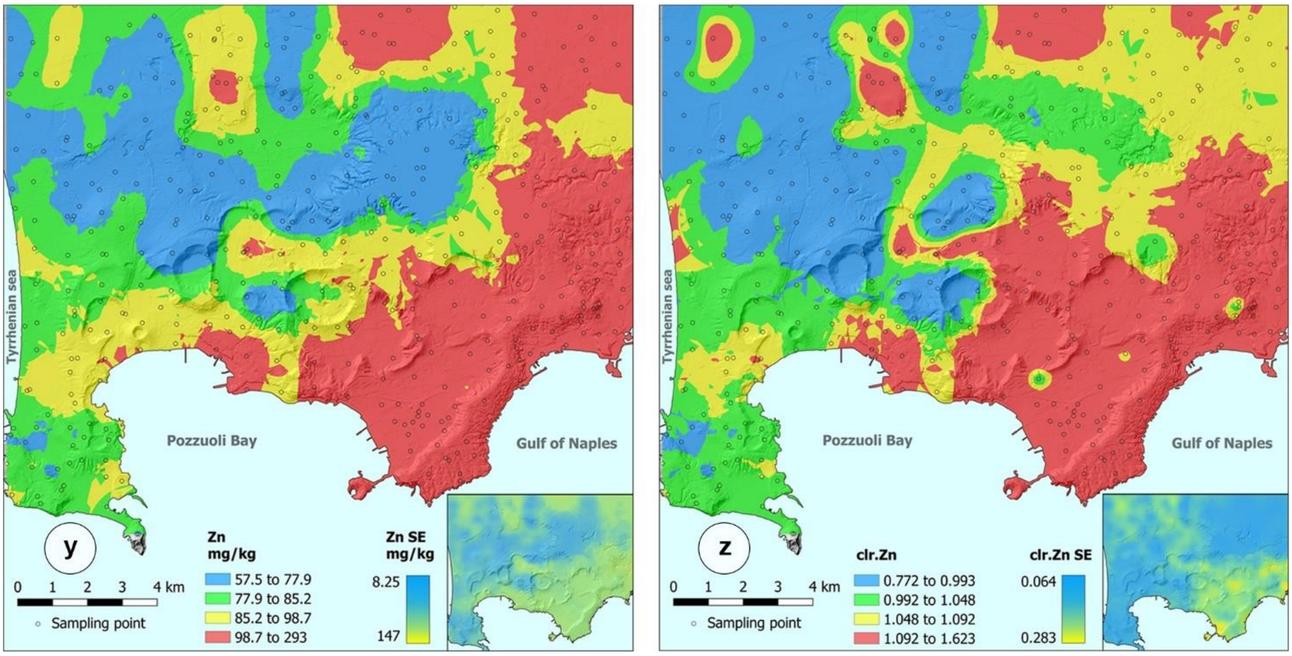


Fig. 2 (continued)

Table 3 The cross-validation prediction errors obtained for the empirical Bayesian kriging model (n = 394) of potentially toxic elements (PTEs) and some other elements carrying important information in the study area. ME is Mean Error; RMSE is Root Mean Square Error; MSE is Mean Standardized Error; RMSSE is Root Mean Square Standardized Error; and ASE is Average Standard Error. The variability of predictions is correctly assessed if ASE ≈ RMSE and RMSS ≈ 1, overestimated if ASE > RMSE and RMSSE < 1, and underestimated if ASE < RMSE and RMSSE > 1 (Kresic and Mikszewski, 2012; Li et al., 2020).

Variable	Element	Transformation	ME	RMSE	MSE	RMSSE	ASE
PTEs	**As**	None	-0.401	11.23	-0.067	1.52	7.29
		Clr-transformation	0.017	0.32	0.030	0.97	0.34
	Cd	None	-0.001	0.42	0.007	1.08	0.35
		Clr-transformation	0.010	0.46	0.025	0.95	0.48
	Co	None	-0.031	2.86	-0.013	1.21	2.12
		Clr-transformation	0.002	0.23	0.014	0.98	0.22
	Cr	None	-0.085	11.82	-0.026	1.27	8.96
		Clr-transformation	0.011	0.51	0.016	0.99	0.52
	Cu	None	-1.454	78.58	-0.024	1.02	76.79
		Clr-transformation	0.002	0.74	0.004	0.98	0.75
	Hg	None	-0.003	0.21	-0.016	1.15	0.18
		Clr-transformation	0.001	0.66	0.004	0.99	0.67
	Ni	None	-0.069	7.44	-0.026	1.38	5.19
		Clr-transformation	0.010	0.37	0.019	0.98	0.38
	Pb	None	-3.838	138.96	-0.005	1.02	102.64
		Clr-transformation	-0.0004	0.42	0.0060	0.95	0.44
	Sb	None	-0.110	2.67	-0.103	1.85	1.54
		Clr-transformation	0.003	0.56	-0.001	0.95	0.58
	Se	None	0.000	0.19	0.004	1.04	0.19
		Clr-transformation	0.008	0.56	0.019	0.99	0.57
	Tl	None	-0.007	0.32	0.014	0.99	0.32
		Clr-transformation	-0.005	0.32	-0.009	0.97	0.34
	V	None	0.353	11.58	0.030	1.11	11.15
		Clr-transformation	0.001	0.23	-0.002	0.99	0.23
	Zn	None	3.424	117.71	0.019	1.01	114.85
		Clr-transformation	0.001	0.36	-0.007	1.01	0.37
Other elements	**Al**	None	102.324	7351.94	0.009	0.98	7484.67
		Clr-transformation	-0.011	0.29	-0.030	0.98	0.30
	K	None	76.794	3754.71	0.014	1.05	3716.69
		Clr-transformation	-0.007	0.36	-0.013	0.98	0.36
	Na	None	40.741	1824.08	0.013	1.05	1922.82
		Clr-transformation	-0.00002	0.43	-0.00154	0.99	0.43
	Th	None	-0.025	3.64	-0.011	1.01	3.59
		Clr-transformation	-0.006	0.39	-0.013	1.00	0.39
	U	None	0.010	1.10	0.001	0.99	1.08
		Clr-transformation	-0.002	0.31	-0.003	1.01	0.31

The Na–K–B association is feasibly indicative of weathering degree in the study area. Phlegrean Fields, Campanian Ignimbrite and Mt. Somma–Vesuvius are the products of potassium-rich volcanism during the Quaternary (Washington, 1906) and previous studies made a connection between the high Na and K concentrations in soil and the youngest volcanic rocks of Phlegrean Fields and Mt. Somma–Vesuvius (De Vivo et al., 2016). In the Campania region, greater Na/K was reported in the soils around Phlegrean Fields, Ischia

Island and Mt. Roccamonfina. The potassic and ultrapotassic rocks together with volcano-sedimentary deposits probably contributed to the sodium level in soil (Thiombane et al., 2019). The relationship between Neapolitan Yellow Tuff caldera and the elevated quantities of potassium and sodium is also confirmed in this study (Fig. 5d and f).

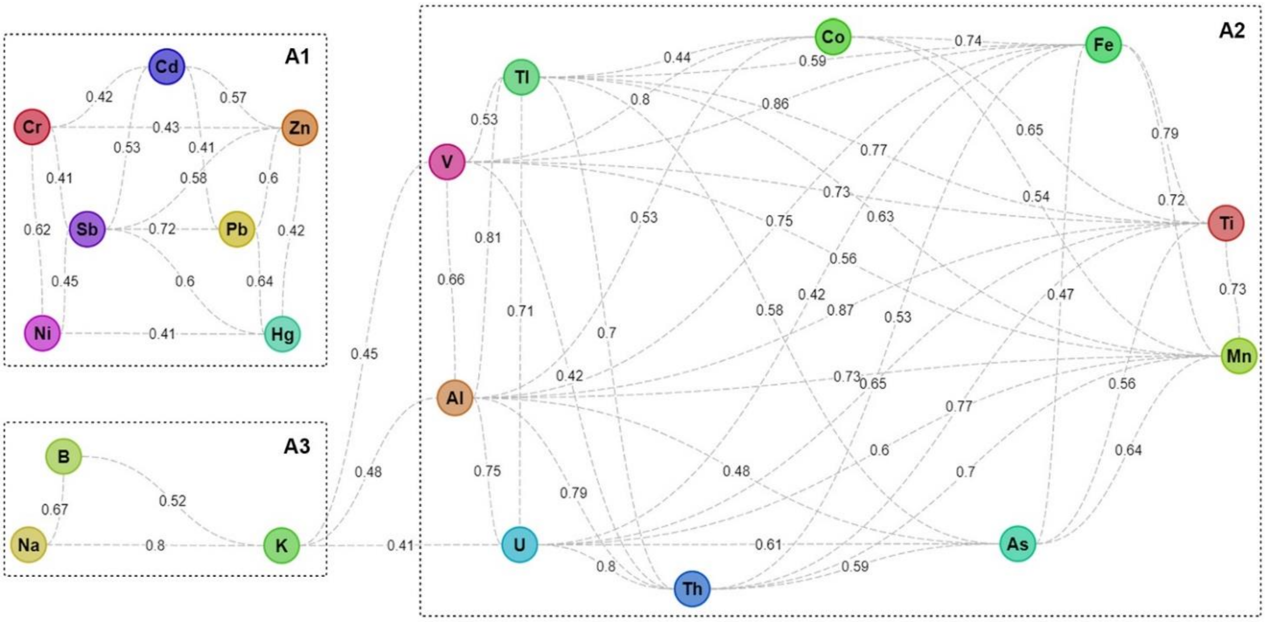


Fig. 3 Correlation diagram of the elements and the probable three associations (i.e. A1 to A3). The Spearman correlation matrix (Table 1) was calculated with symmetric coordinates and a threshold of 0.4 was considered for generating the diagram. Se and Cu do not exist in the diagram because of their low correlation coefficients.

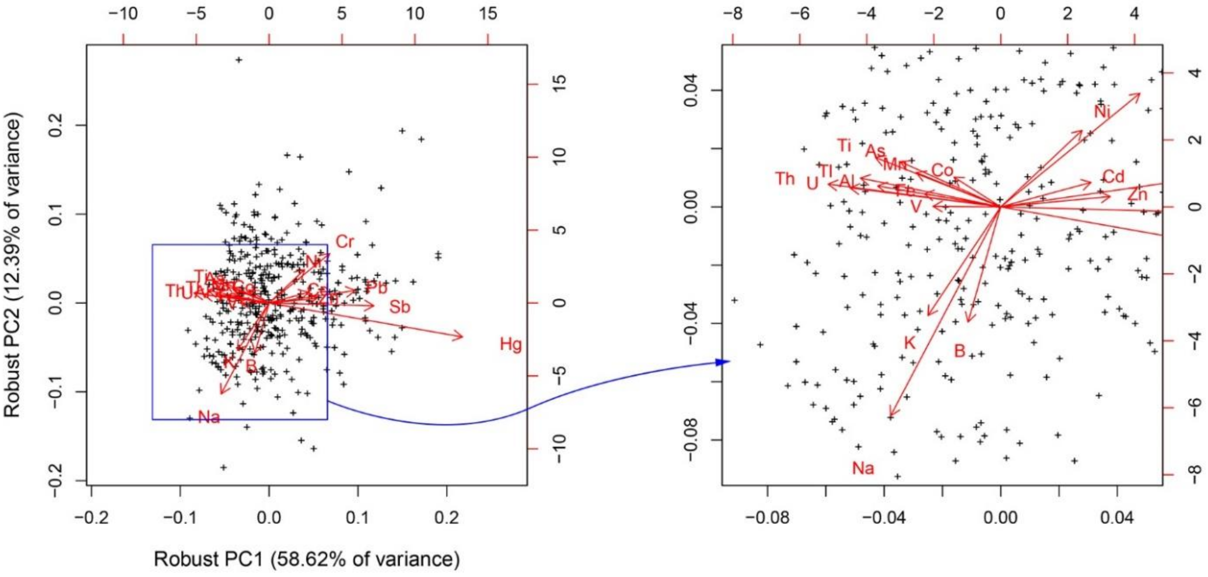


Fig. 4 Covariance biplot of the data projected on the first two principal axes (explained variance: 71%). The right panel is the zoomed area of the blue box. Input variables are the same as the correlation diagram (Fig. 3).

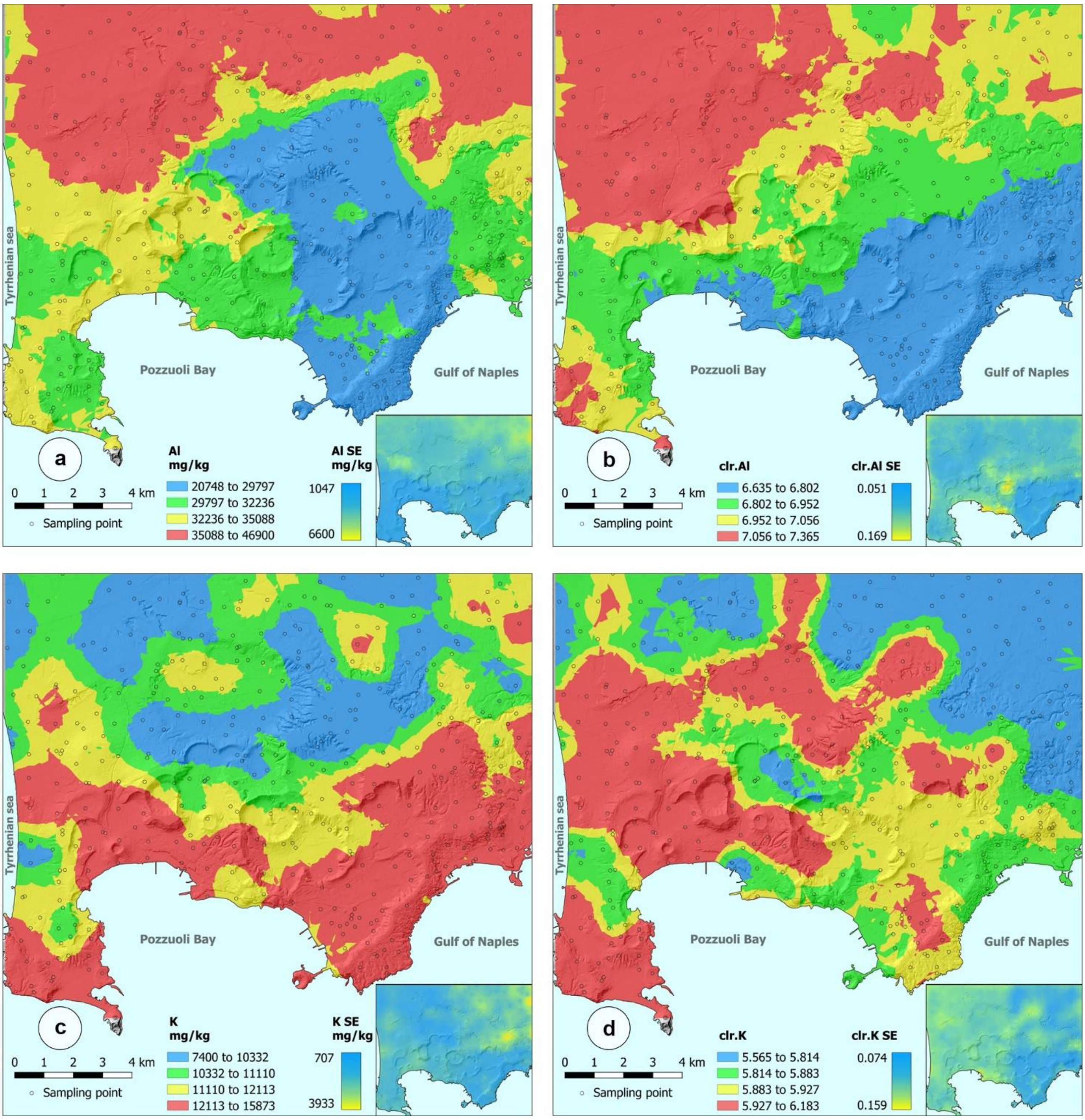


Fig. 5 Spatial distribution of the elements that help to clarify the geochemical processes in soil: Al (a and b), K (c and d), Na (e and f), Th (g and h) and U (i and j). The left and right geochemical maps are generated by raw and clr-transformed values, respectively. The lower right insets are the prediction standard error maps.

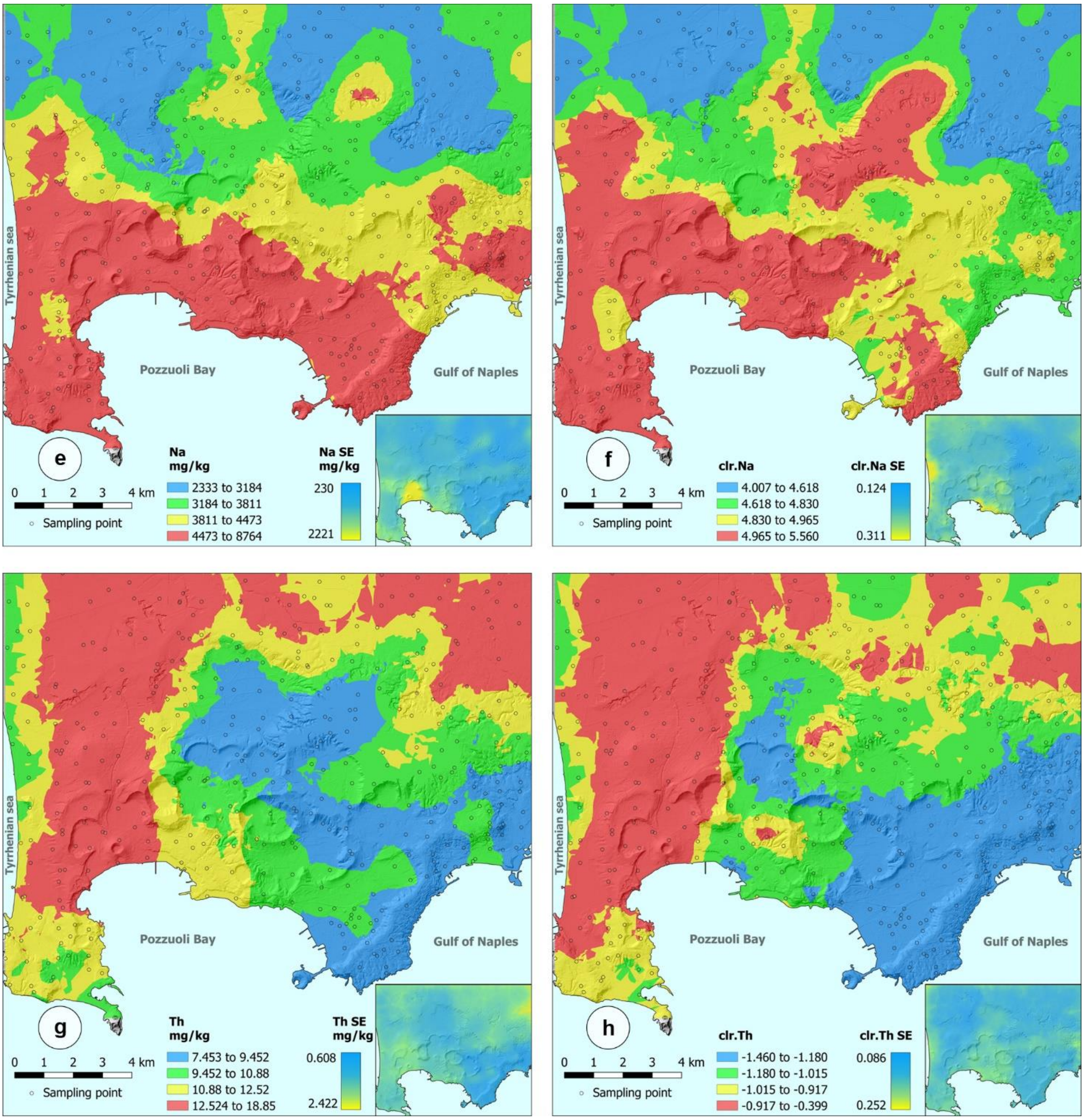


Fig. 5 (continued)

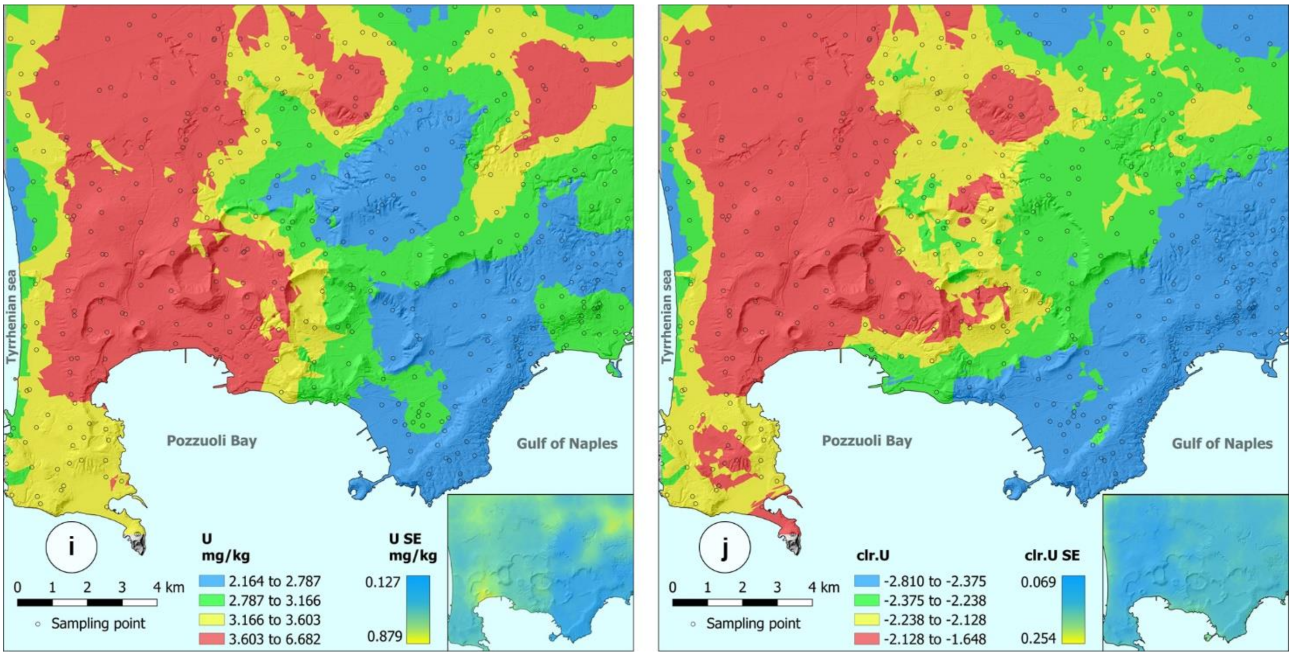


Fig. 5 (continued)

The soil samples are categorized into two clusters on the basis of either PC1 or PC2 scores. The PC1 scores distinguish the anthropogenic vs. geogenic impact while the PC2 scores differentiate the less weathered from the more weathered soils (Fig. 6). Evidently, the soil is under higher anthropogenic pressure in Naples urban area (Fig. 6a) which is in agreement with higher population density in the eastern sector. The soil composition is much more similar to the pyroclastic deposits in west. Regarding Fig. 6b, the less weathered soil samples are mostly located inside the NYT caldera, but the weathering processes become more intensive outside the caldera rim. To verify these hypotheses, the Pb–U–Th and Al–Na–K three-part subcompositions are extracted from the soil dataset and plotted with the literature data on rock composition in ternary diagrams (Fig. 7). Composition of the soil predominantly derived from geogenic origin shares similarities with that of post-15 ka tephra. However, the concentration of lead increases in the soil impacted by various anthropogenic activities in the urban environment (Fig. 7a). Aluminum, sodium and potassium quantities of tephra and less weathered soil are roughly the same, but concentrations of Na and K decrease with increasing weathering intensity (Fig. 7b). The geochemical maps of Pb, U, Th, Al, Na and K are also pieces of supporting evidence (Figs. 2 and 5).

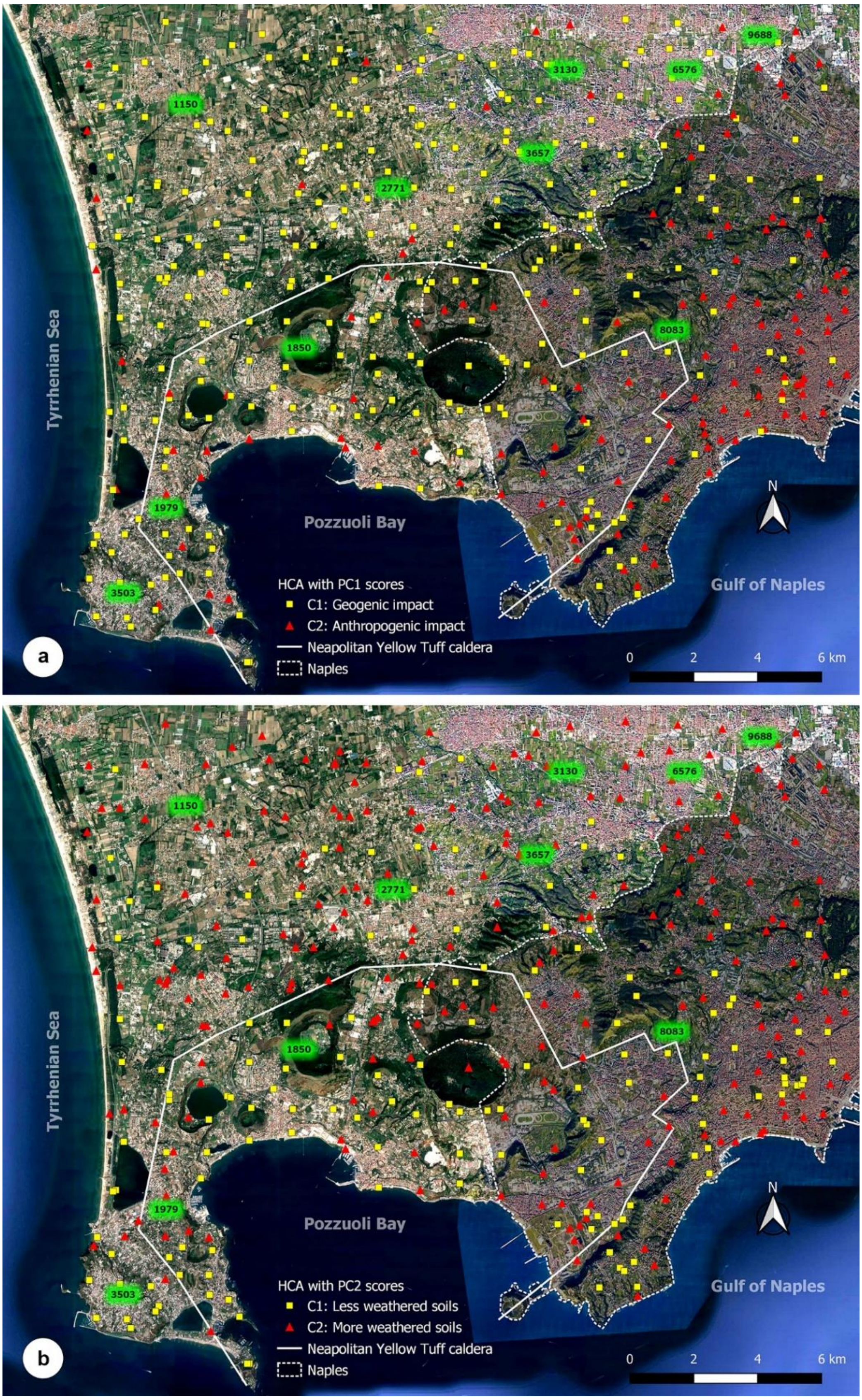


Fig. 6 The spatial distribution of clusters for (a) PC1 and (b) PC2 scores. C1 and C2 refer to the first and second clusters identified after computing the Euclidean distances between the observations and applying the Ward's method for agglomerative clustering. The numbers with green halos represent population density. The Neapolitan Yellow Tuff caldera rim was proposed by Vitale and Isaia (2014).

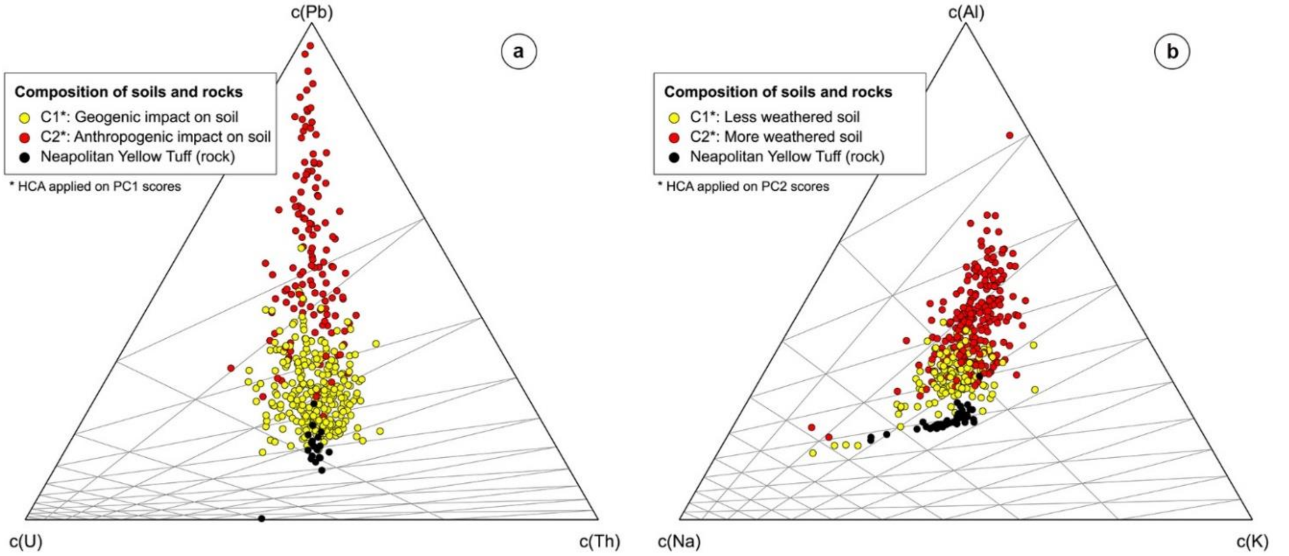


Fig. 7 A comparison between geochemical composition of the post-15 ka tephra from Phlegrean Fields (Smith et al., 2011) and the soil samples in this study: (a) ternary diagram of the subcomposition Pb–Th–U, and (b) ternary diagram of the subcomposition Al–K–Na. The data are centered to obtain a better view of the pattern. Spatial distribution of each cluster in the left and right panels is shown in Fig. 6.

Table 4 Percentage of the soil samples exceeding the PTEs contamination thresholds for residential use (CSC$_R$) and for industrial/commercial use (CSC$_I$). All values are in mg/kg.

PTE	**CSC$_R$ ^{a}**	**% above CSC$_R$**	**CSC$_I$ ^{a}**	**% above CSC$_I$**
Cu	120	24.6	600	0
Pb	100	23.1	1000	0.8
Zn	150	19.5	1500	0
Ni	120	0	500	0
Co	20	0.8	250	0
As	20	8.1	50	1.8
Cd	2	1.0	15	0
Sb	10	1.3	30	0.3
V	90	1.0	250	0
Cr	150	0	800	0
Tl	1	74.4	10	0
Hg	1	1.3	5	0
Se	3	0	15	0

[a] D.Lgs.152/2006

4.3. Health risk for the local children

To choose the PTEs with potential health risks for the locals, the PTE concentrations in soil samples are compared with the corresponding contamination thresholds for different land uses (i.e. residential and industrial/commercial). Because copper and selenium are not considered for statistical analysis and source identification, they are excluded in health risk assessment as well. According to Table 4, Tl, Pb and Zn contents exceed the threshold for residential use in more than 15% of the collected samples. Thallium which comes from a natural source is also culled before health risk evaluation. Regarding Figs. 2 and 6, the elevated Pb and Zn quantities mainly correspond to the urbanized areas with high population density. It is noteworthy that the

PTE levels are barely greater than the thresholds established by the Italian legislation for industrial/commercial use of soil.

In this chapter, children (<6 years old) are considered for health risk assessment because: (1) Pb has particularly significant adverse health effects in children; and (2) children under 6 years old have a higher risk of exposure due to more frequent hand-to-mouth behaviour (USEPA, 2004). Therefore, children are categorized into six age brackets (i.e. 0-1, 1-2, 2-3, 3-4, 4-5 and 5-6 years old) and a probabilistic approach is implemented to separately evaluate the non-carcinogenic and carcinogenic health risks for each category due to the physiological differences. The probability distributions and their parameters for each age range are listed in Table 5. The reference doses and cancer slope factors are also shown in Table 6.

Table 6 The oral reference dose (RfD), inhalation reference concentration (RfC) and cancer slope factor (CSF) for lead and zinc. All values are in mg/kg day^{-1}, except for RfC_{inh} (mg/m^3 day^{-1}). The data are retrieved from the Risk Assessment Information System (2021) and Masri et al. (2021).

PTE	**RfD_{ing}**	**RfC_{inh}**	**CSF_{ing}**	**CSF_{inh}**
Pb	0.0035	-	0.0085	1.20E-05
Zn	3.00E-01	-	-	-

The results of health risk assessment for children are visualized in Fig. 8. The non-carcinogenic health risk (HQ > 1) is characterized for exposure of all age ranges to Pb in the soil through ingestion route (Fig. 8a). However, the certainty does not exceed 5% (Table 7). The hazard quotients of Zn are far below one which is not indicative of a non-carcinogenic health threat to children due to soil ingestion (Fig. 8b). The obtained CR_{ing} values for Pb are under the threshold of 1E-06 and highlight negligible carcinogenic health risk (Fig. 8c). In stark contrast, the CR_{inh} of lead surpasses 1E-06 to some extent and poses an acceptable health risk to the children in all age groups (Fig. 8d). The certainty of exceeding this threshold mainly ranges from about 25% in the 5-6 years old children to approximately 45% in the 2-3 years old individuals. For the children aged 0 to 1 year, the certainty value is the highest (i.e. 90%).

In the Naples urban area, Imperato et al. (2003) revealed the increased Pb and Zn levels in the surficial soil layer (<2 cm depth) between 1974 and 1999, particularly along the roads. Pizzolante et al. (2021) reported that Naples is one of the municipalities under the highest environmental pressure in the Campania Region and generated a regional-scale map of environmental pressure index for each municipality to encourage planning remediation programs, human biomonitoring studies and public health interventions. Although the results of this investigation do not show noticeable health risk in the study area, it should not be overlooked because Pb is one of the leading persistent pollutants across the world that can undergo bioaccumulation processes and negatively affect humans, animals and ecosystems (Siegel, 2002). A range of health issues are generally linked to the excess lead concentrations, including damage to the nervous systems, anemia, inhibition of heme formation, kidney damage, impaired mental development in young children, genotoxicity, impaired reproductivity and carcinogenicity (Kabata-Pendias and Mukherjee, 2007).

Table 5 The probability distribution of each variable for considering uncertainty in health risk assessment (Haddad et al., 2001; USEPA, 2011; ISTAT, 2019). S.D. refers to standard deviation.

Age (year)	**Body weight (kg)**		**Ingestion rate (mg/day)**		**Inhalation rate (m^3/day)**		**Exposure frequency (day)**		**Life expectancy (year)**	
	Distribution	**Parameter**	**Distribution**	**Parameter**	**Distribution**	**Parameter**	**Distribution**	**Parameter**	**Distribution**	**Parameter**
0 to 1	Lognormal	Min = 2.8 Mean = 6.54 S.D. = 2.83	Triangular	Min = 1 Mode = 30 Max = 50	Triangular	Min = 3.5 Mode = 5.4 Max = 12.8	Triangular	Min = 15 Mode = 350 Max = 365	Uniform	Min = 79.6 Max = 83.8
1 to < 2	Uniform	Min = 8.4 Max = 11.5	Uniform	Min = 50 Max = 200	Uniform	Min = 5.4 Max = 9.2				
2 to < 3	Uniform	Min = 10.7 Max = 14.5				Min = 8.9 Max = 13.7				
3 to < 4	Uniform	Min = 12.5 Max = 16.7				Min = 10.1 Max = 13.8				
4 to < 5	Uniform	Min = 14.2 Max = 19.7								
5 to < 6	Uniform	Min = 18.1 Max = 25.6								

Table 7 The non-carcinogenic and carcinogenic health risk for exposure to lead and zinc through ingestion and inhalation routes based on Monte Carlo simulation. HQ: hazard quotient, CR: cancer risk, S.D.: standard deviation. HQ > 1 shows potential adverse health effects, $10^{-4} > CR > 10^{-6}$ denotes the acceptable risk and $CR \geq 10^{-4}$ represents the unacceptable risk (USEPA, 2011). The proportions show the certainty level for exceeding a given threshold based on 50000 simulations.

	PTE	**Descriptive statistics**	**Age (year)**					
			0 to 1	**1 to 2**	**2 to 3**	**3 to 4**	**4 to 5**	**5 to 6**
HQ for the ingestion route	Pb	Min	1.00E-04	1.40E-03	1.20E-03	1.80E-03	1.00E-03	9.00E-04
		Mean	8.73E-02	2.34E-01	1.88E-01	1.61E-01	1.38E-01	1.07E-01
		Median	4.21E-02	1.18E-01	9.40E-02	8.07E-02	6.98E-02	5.40E-02
		Max	1.23E+01	1.63E+01	1.85E+01	1.04E+01	9.92E+00	8.21E+00
		S.D.	1.70E-01	4.15E-01	3.41E-01	2.83E-01	2.44E-01	1.95E-01
		$HQ_{ing} > 1$	0.52%	3.44%	2.18%	1.69%	1.24%	0.67%
	Zn	Min	8.38E-06	3.69E-05	3.50E-05	3.21E-05	3.29E-05	2.90E-05
		Mean	1.20E-03	2.70E-03	2.70E-03	2.20E-03	2.00E-03	1.50E-03
		Median	8.00E-04	1.90E-03	1.90E-03	1.60E-03	1.40E-03	1.10E-03
		Max	3.44E-02	8.21E-02	1.36E-01	7.50E-02	5.33E-02	6.63E-02
		S.D.	1.50E-03	2.90E-03	2.90E-03	2.40E-03	2.10E-03	1.70E-03
		$HQ_{ing} > 1$	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
CR for the ingestion and inhalation routes	Pb	Min	8.93E-11	4.75E-10	5.27E-10	3.66E-10	4.22E-10	2.93E-10
		Mean	3.19E-08	8.69E-08	6.76E-08	5.82E-08	5.05E-08	3.94E-08
		Median	1.51E-08	4.33E-08	3.41E-08	2.92E-08	2.53E-08	1.97E-08
		Max	2.25E-06	1.10E-05	1.16E-05	6.51E-06	3.54E-06	9.96E-06
		S.D.	6.08E-08	1.63E+07	1.33E-07	1.03E-07	9.19E-08	8.23E-08
		$10^{-4} > CR_{ing} > 10^{-6}$	0.04%	0.43%	0.22%	0.15%	0.12%	0.04%
		$CR_{ing} \geq 10^{-4}$	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
		Min	2.81E-08	1.73E-08	2.21E-08	2.01E-08	1.48E-08	1.06E-08
		Mean	8.74E-06	1.33E-06	1.66E-06	1.52E-06	1.30E-06	1.01E-06
		Median	4.10E-06	7.05-07	8.69E-07	8.01E-07	6.83E-07	5.34E-07
		Max	1.47E-03	9.49E-05	1.05E-04	7.80E-05	7.19E-05	9.13E-05
		S.D.	1.84E-05	2.19E-06	2.78E-06	2.49E-06	2.14E-08	1.78E-06
		$10^{-4} > CR_{inh} > 10^{-6}$	90.13%	36.84%	44.54%	41.34%	35.90%	27.60%
		$CR_{inh} \geq 10^{-4}$	0.50%	0.00%	0.00%	0.00%	0.00%	0.00%

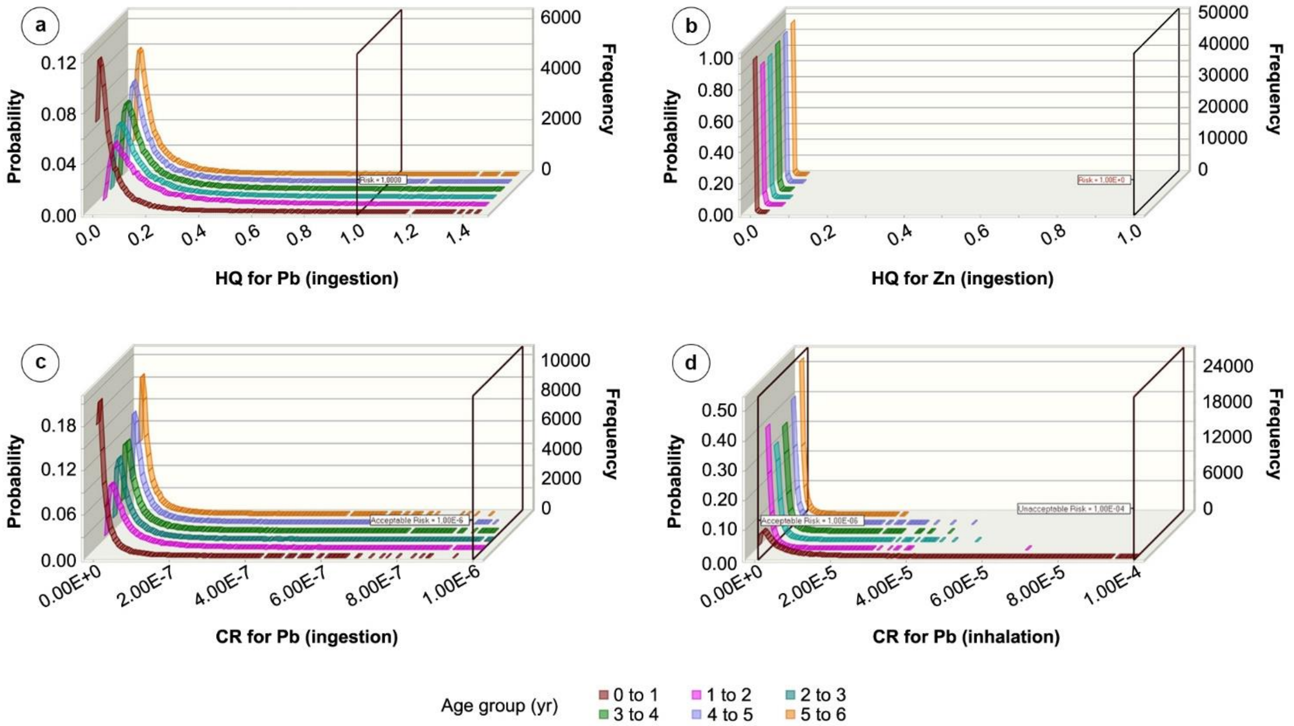


Fig. 8 The non-carcinogenic Pb and Zn (a and b, respectively) health risk together with carcinogenic Pb (c and d) health risk for different age groups of children. A total of 50000 trials were implemented in the Monte Carlo simulation.

5. Conclusions

Recently, the global population in urban areas exceeded that of rural areas which could negatively impact quality of life, particularly when there are official travel restrictions such as the ongoing COVID-19 outbreak. Naples is one of the highly populated metropolises in Italy that partially occupies Phlegrean Fields and the immediate surroundings. In addition, Phlegrean Fields is located in a zone that was declared as a contaminated site of national interest due to industrial activities, agricultural practices and buffalo farms in this area. Therefore, 394 soil samples are collected from 0-20 cm depth, passed from a 2-mm sieve and concentrations of PTEs together with some other elements that may help characterize the source are determined. The empirical Bayesian kriging interpolation method, compositional data analysis and probabilistic health risk assessment are then used to achieve the objectives of this chapter. The correlation diagram and robust principal component analysis demonstrate three groups of elements. The Pb–Zn–Hg–Cd–Sb–Cr–Ni association probably highlights the anthropogenic origin (e.g., the heavy traffic load and fossil fuel combustion). The geochemical maps link higher concentrations of these PTEs with Naples urban area in the east. However, the Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association likely reveals the immobile elements during weathering of the pyroclastic deposits from Phlegrean Fields, Roccamonfina and Somma–Vesuvius volcanic eruptions. Contrary to the first association, the relationship between these elements is a stronger in the western part of the study area. Arsenic concentration in west Phlegrean Fields could be attributed to the geothermal activity in the Mofete fields as well. The Na–K–B association is probably indicative of weathering Phlegrean Fields, Campanian Ignimbrite and Mt. Somma–Vesuvius potassium-rich volcanic products. The soil samples with relatively higher Na and K (i.e. less weathering intensity) are spatially related to the Neapolitan Yellow Tuff caldera. For all age groups of children under 6, negligible non-carcinogenic health risk (certainty > 95%) is expected because of exposure to Pb and Zn in soil through ingestion. The CR_{ing} values for Pb are below 1E-06 and highlight negligible cancer risk as well. Nevertheless, exposure to lead through the inhalation pathway poses acceptable cancer risk to children with certainty of about 25% (in the 5-6 years old children) to 45% (in the 2-3 years old individuals). The certainty value is the highest (i.e. 90%) for the children aged 0 to 1 year. Previous studies documented the increased Pb and Zn levels in soil over 26 years which underlines the need to monitor anthropogenic sources of pollution for encouraging sustainability in the community and avoiding future health issues. The results indicate that the recent advances are advantageous in geostatistical mapping, characterizing origin of PTEs and assessing the associated health risk.

Contribution of the Ph.D. candidate

Determining research objectives; Data visualization; Data interpretation; Writing the original draft; Revising the draft

CHAPTER 6

Scientific Collaborations

Section 6.1 was published in "Journal of Geochemical Exploration" as a Research Article:

Zheng, C., Luo, X. *, Wen, M. *, Ebrahimi, P., Liu, P., Liu, G., ... & Albanese, S. (2020). Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China. Journal of Geochemical Exploration, 213, 106509. https://doi.org/10.1016/j.gexplo.2020.106509.

* Corresponding authors: lxr811@glut.edu.cn (X. Luo); meilanwen112@126.com (M. Wen).

Section 6.2 was published in "Geosciences" as a Review Article:

Ebrahimi, P., & Barbieri, M. * (2019). Gadolinium as an emerging microcontaminant in water resources: threats and opportunities. Geosciences, 9(2), 93. https://doi.org/10.3390/geosciences9020093.

* Corresponding author: maurizio.barbieri@uniroma1.it

Section 6.3 was submitted to a journal as a Research Article:

Guarino, A., Albanese, S. *, Cicchella, D., Ebrahimi, P., Dominech, S., Allocca, C., Romano, N., De Vivo, B., & Lima, A. Selected major and potentially toxic elements bioavailability in agricultural soils of Campania region (Italy): The spatial patterns and influencing factors – *Submitted*

* Corresponding author: stefano.albanese@unina.it

Section 6.1:
Deep Orebody Prediction in the Ashele Cu-Zn Deposit, NW China

Abstract

In this investigation, a total of 404 rock and ore samples were systematically collected from 5 boreholes in two exploration lines of the Ashele copper-zinc polymetallic mining area in Xinjiang. Twelve elements were analyzed to evaluate the potential for deep and periphery prospecting. Subsequently, the concentration zoning and the vertical element zoning sequence of the primary halo were graphically generated. Based on both previous studies and the result of the actual analysis, it was determined that As, Sb, Ba, and Au are the supra-ore halo elements, Pb, Zn, Cu, and Ag (Au) are the near-ore halo elements, and Bi, Co, Mo, and Sn are the sub-ore halo elements. Factor analysis revealed exhalative-sedimentation mineralization as the first metallogenic stage followed by metamorphic-deformation and hydrothermal superimposed mineralization as the second one. The superposition area of the sub-ore and supra-ore halos and the multivariate prospecting information (including the primary halo zoning index, the geochemical parameter, the trilinear diagram of the primary halo and the factor score map) suggest that there may be blind orebodies in the deep parts of four boreholes. Some geochemical parameters revealed the multiple-periodic and multiple-staged overlapping mineralization signatures.

1. Introduction

Primary halos are created by the interactions between the ore-forming fluids and the surrounding rocks (Parsapoor et al., 2017; Gong et al., 2016). The characteristics and scale of the halo zonation are related to the concentrations of elements in the ore-bearing solutions, the relative mobility of elements, the background contents of elements in the enclosing rocks and the mineralization type (Beus and Grigorian, 1977). Primary halos may be hundreds of times larger than the deeply buried deposits (Hosseini-Dinani and Aftabi, 2016; Helba et al., 2020) and appear on the ground surface (Liu et al., 2010). The axial zonation is in accord with the vertical zonation in steeply dipping orebodies and expresses flow direction of the ore-bearing solutions whereas the transversal zonation reflects mineralization across the orebody (Beus and Grigorian, 1977; Chen and Zhao, 1998; Luo et al., 2007; Shao, 1997).

The Ashele deposit is a typical VMS type deposit located in the Ashele basin on the southwest margin of the Altai orogenic belt in China. The main metallic minerals are Cu and Zn associated with Au and Ag. Detailed compositional, structural, genetic and geological investigations were carried out systematically (Liu et al., 2010; Wu et al., 2015), but primary halo of the deposit has not been systematically studied yet. This work aims to provide the basis for prospecting prediction of the deep and peripheral sections of the mining area.

2. Deposit geology

The Ashele formation is divided into two lithologic groups (Fig. 1): D_2as^1 which is less distributed in southeast, dominated by tuff and intercalated with crystalline limestone; and D_2as^2 which is in conformable contact with the first lithologic group and exposed in the middle and northwest of the mining area. In this study, the Ashele No. 1 copper-zinc deposit is of interest which occurs in the middle layer (D_2as^{2b}) of the second group. It is distributed in the S-N direction, lies between basalt and tuff, and integrates with the strata. Elevation of the orebody ranges from -600 to 888 m and the thickness varies between 5 and 120 m (Yang et al., 2016). The massive sulfide Cu-Zn deposit, dominated by volcanic exhalation-sedimentary genesis and superimposed by the metamorphism and hydrothermal fluids in the later stage, shows the common characteristics of a typical submarine volcanic-exhalative-sedimentary deposit. The ore metal minerals are mostly pyrite, chalcopyrite and sphalerite, followed by tetrahedrite, galena and other metal sulfide minerals. Gangue minerals are mainly quartz, sericite (muscovite) and contain small amounts of chlorite, barite, calcite, dolomite, rutile and sphene. The structure of ore is mainly massive and strip structures, followed by disseminated-banded and veinlets-disseminated structures with a tiny amount of breccia structure.

3. Materials and methods

A total of 404 rocks were collected from Zk2102, Zk2103, Zk2107, Zk2504 and Zk2505 boreholes in the No. 21 and No. 25 cross-sections (Fig. 1) and concentrations of As, Sb, Bi, Au, Ag, Pb, Mo, Sn, Cu, Zn, Ba and Co were determined. Based on replicate measurements, the precision (RSD%) of chemical analysis ranges between 2 and 5%.

To obtain the zonality sequence of halo elements, an improved version of the Grigorian method was used as explained in Wang et al. (2007). The distribution maps of axial concentration zoning for the primary halo were also visualized through the inverse distance weighting (IDW) method. The concentration-area plot (Cheng et al., 1994) was then applied for determination of three intervals (i.e. the outer, middle and inner bands, respectively).

An R-mode factor analysis (FA) was also performed after applying a log-transformation to categorize variables into a few factors regarding the correlation matrix (Field, 2013; Reimann et al., 2008). Barium and Co were excluded because of low correlation (<0.6) with the other elements. In addition, Sb was eliminated due to low communality (<0.7). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO = 0.901) revealed meeting the requirements for FA. After a varimax rotation with Kaiser Normalization, two components were extracted by means of the principal component analysis method.

Finally, Vertical Element Contrast Coefficient (VECC) and Geochemical Parameter (GP) that intuitively demonstrate vertical distribution of a signal element or an elemental association were implemented. Larger vertical GP values represent greater possibility of finding a concealed orebody at depth (Zhang et al., 2010). A trilinear diagram of the primary halo was also plotted for each borehole using the normalized dataset to illustrates the association between orebody and mineralized alternation zone in different elevation segments. The supra-ore (As and Sb), near-ore (Cu, Pb and Zn) and sub-ore (Bi, Mo and Sn) halo elements were considered as three separate groups and the value of each diagram axis was defined by the arithmetic sum of each group.

4. Result and discussion

4.1. Primary halo axial-zoning sequence

The primary halo zoning during magma migration is caused by the gradual evolution of the metallogenic fluid properties and the conditions generating a range of mineral assemblages (Shao, 1997). Based on the dataset in the present study, the supra-ore halo elements are As, Sb, Ba, and Au, the near-ore halo element assemblage is Pb, Zn, Cu, and Ag (Au), and the sub-ore halo elements are Bi, Co, Mo, and Sn in the Ashele deposit. All boreholes indicate at least three groups of primary halo zoning. Although quantity of the primary halo anomalies can be related to orebodies and/or mineralized alteration zones, some of them still remain unexplained. A superposition of the sub-ore and supra-ore halos is observed at the end of Zk2102, Zk2103 and Zk2505 boreholes, revealing a series of small known orebodies beneath the main one. Besides, some anomalies in Zk2504, Zk2107 and Zk2505 boreholes do not correspond to the orebody or alteration zone, indicating the probable concealed blind ore bodies.

4.2. Primary halo concentration zoning

Regarding the axial concentration zoning of the primary halo in the No. 21 exploration line, the elements (except for Ba and Co) develop a clear inner anomaly zone corresponding to the ore deposits while the outer and middle zone anomalies are not clearly explainable. Copper and Zn (Pb) are the main metallogenic and near-ore halo elements, whose anomalies are mostly related to the orebodies. Because Sb and As are the supra-ore halo elements, they show good abnormal continuities in the whole axial zoning and the anomaly scales are wider than those of Cu and Zn, reflecting their stronger diffusivity and higher mobility.

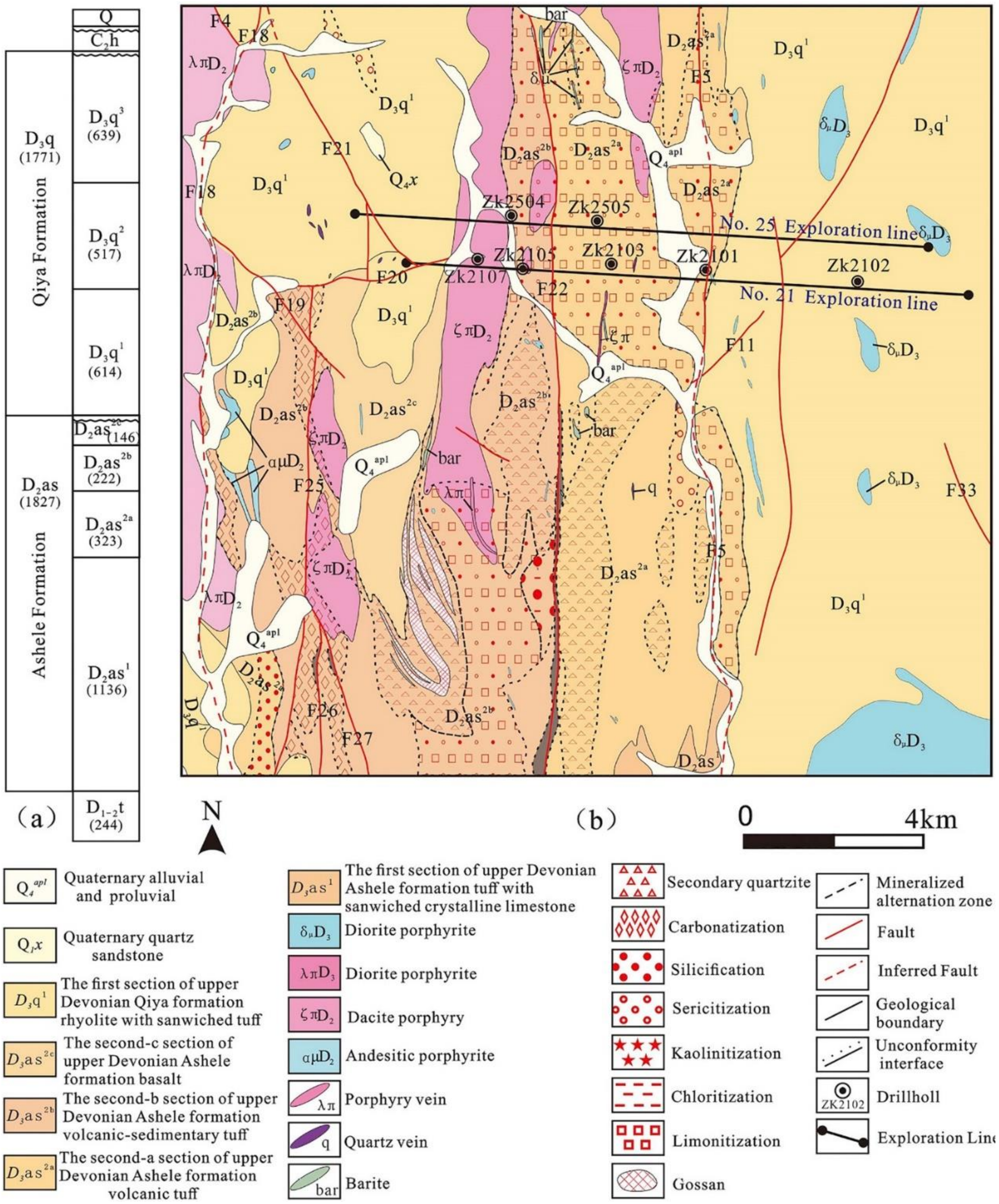


Fig. 1 (a) Stratum histogram of the study area, (b) geologic map of the copper-zinc deposit.

In the axial concentration zoning of the primary halo in the No. 25 exploration line, a fracture zone is extended from the surficial part of Zk2504 borehole towards depth and intercepts the Zk2505 borehole at an elevation of 700 m (Fig. 1) which is likely responsible for the significant anomaly of most elements (e.g., Cu, Zn and Pb) at elevations 600-900 m. Furthermore, the small-scaled inner Pb anomalies occur between -100 and 50 m and there is limitated development of Cu anomalies at depth. Copper and Zn along with Mo and Ag are characterized by a series of inner anomalies at 300-500 m that may have been caused by intrusive rocks. Arsenic, Sb and Ba are the supra-ore halo elements that show remarkable anomalies near the surface (700-900 m) although they may have been subject to near-surface weathering and denudation. Moreover, these elements have good inner zone anomalies at levels from -100 to 50 m and reveal a good spatial correspondence with the "I-1" and "I-2" orebodies.

4.3. Factor analysis

About 80% of the total variance is explained in factor analysis (Table 1) and both factors have a good association with the mineralization processes. F1 refers to the high-temperature indicator elements (Mo, Sn and Bi), metallogenic elements (Cu and Ag) and low-temperature indicator elements (As). However, F2 is related to a subset of metallogenic elements including Zn, Pb, Au and Ag. Previous studies (Chen, 1996; Gao et al., 2010; Wang, 1995; Wu et al., 2015; Ye et al., 1998) showed that after formation of the main layered orebodies during the exhalative-sedimentation period, the Ashele copper-zinc deposit undergone multistage reformations such as metamorphism, magmatic-hydrothermal superposition and epigenetic oxidation. Concerning the extracted factors (Table 1) and the temperature of ore-forming fluids in the mine range, the F1 association is mostly related to metamorphism and hydrothermal superposition, but the F2 assemblage mainly explains the mineralization of exhalative-sedimentation deposits. Furthermore, both factors contain metallogenic elements (e.g., F1: Cu (Ag); and F2: Zn (Pb)) which likely show the multistage superposition genesis of the deposit.

Table 1 Varimax-rotated factor analysis of the selected variables for the rock samples of both exploration lines.

Variables	Factor		Communality
	F1	F2	
As	**0.647**	0.594	0.772
Bi	**0.899**	0.239	0.865
Sn	**0.903**	0.135	0.834
Ag	**0.657**	**0.614**	0.809
Pb	0.397	**0.825**	0.838
Mo	**0.793**	0.355	0.755
Cu	**0.740**	0.399	0.707
Zn	0.077	**0.867**	0.757
Au	0.584	**0.689**	0.816
Eigenvalues	68.227	11.248	
Total variance (%)	46.174	33.301	
Cum. total variance (%)	46.174	79.475	

Bold entries: loading values over 0.60.

4.4. Vertical variation of primary halo characteristics

The VECC and GP of both cross-sections well reveal the orebody spatial distribution. Enrichment of the supra-ore halo directly indicates the deep copper-sulfur industrial orebody at the tail of Zk2103, Zk2102 and Zk2505 boreholes. The Zk2107 and Zk2504 (Zk2102 and Zk2505) also show deep concealed orebodies or mineralized alteration zones to a certain extent. An orebody and a mineralized alteration zone can result in a high VECC of supra-ore halo and an increasing tendency in GP, but the intensity of VECC and GP for the orebody is stronger than that of mineralized alteration.

4.5. Trilinear diagram of the primary halo

Trilinear diagrams were generated to better investigate the relationship between different ore halos and areal distribution of orebody or mineralization zone. The development of sub-ore halo in Zk2102 is relatively high although the near-ore halo is more evident than the supra- and sub-ore halos in Zk2103. Most of the Zk2107 and Zk2505 samples fall in the near-ore halo area and Zk2504 indicates better development of the supra-ore halo. Clearly, the mineralized alteration zones and orebodies are well distributed in the sub-ore halo area of each borehole in the No. 21 exploration line (Fig. 1), highlighting a powerful indicator for prospecting. The segments at the end of the boreholes fall into the overlapping area (superposition of supra- and sub-ore halos) which suggest high possibility for the existence of a deep orebody or mineralized alteration zone. It is consistent with the judgment criterion of the superimposed halo of the blind orebody structure proposed by Li et al. (1995, 2016), Li (1998) and Li and Zhang (1999).

4.6. Primary halo superposition model and deep target delineation

On the basis of the previous discussion, the ideal model for the superposition of primary halo is constructed in Figs. 2 and 3 and four prospecting targets are identified. The first one is the prediction area (a) in Fig. 2, where the small “I-1” orebody between -300 and 100 m is attached to the “I-2” orebody at the -100 to 400 m elevation interval. Scale of the primary halo concentration for most elements in this section (from -300 to 0 m) is comparable to that of the upper section (between 0 and 400 m), but larger than that of the upper anomaly (e.g. Zn, Cu, Au, As, Sb, Sn, Mo and Bi). In addition, the prediction area (a) with a superposition of sub-ore and supra-ore halos at the tail of Zk2107 borehole probably indicates a hidden orebody. Considering the areal distribution of the known orebodies (lenticular or stringer assemblages and parallel or attach to each other) in Fig. 2, it could be inferred that hidden parallel orebodies with the shape of an en echelon might exist in the prediction area (a). Furthermore, the lithology of this area is consistent with that of the ore-bearing strata.

The prediction area (b-1) (Fig. 2) occurs in favorable ore-bearing strata and develops a superposition of sub- and supra-ore halos along with the evident primary halo anomalies of many elements, which can be possibly caused either by the “I-3” and “I-4” orebodies or a hidden orebody. According to the increasing GP, there is high possibility of a hidden orebody or mineralization zone beneath the “I-3” and “I-4” orebodies. The same favorable conditions mentioned above can also be adapted to the No. 25 cross-section (Fig. 3) at the elevations from -400 to -600 m. Regarding the extension of orebody from south to north, there is slight possibility that the “I-3” and “I-4” orebodies extend towards the (b-2) area.

The prediction area (c) (Fig. 3) shows significant primary halo concentrations of Bi, Cu, Pb and As, the high VECC values at 650 m and the fluctuating downtrend of GP between 600 and 800 m. On the other hand, there is a strong association between Zn, Pb, Au and Ag in the area (c). Furthermore, the area (c) is adjacent to the fracture zone and the lithology is dacite porphyry, possibly corresponding to the mineralized alternation zone. The prediction area (d) shares similarities with area (c) in the lithology, the superimposed sub- and supra-ore halos along with the abnormal vertical geochemical values. The small fracture zone in dacite porphyry on the lower right side of the prediction area (d) can explain these anomalies.

5. Conclusions

The dataset obtained from 404 rock samples demonstrates that the axial primary halo methods and statistical techniques are advantageous for deep copper-zinc orebody prediction in the Ashele mining area. The characteristics of superimposed haloes for ore deposit are as follows:

- The supra-ore halo element assemblage is As–Sb–Ba–Au, the near-ore halo element assemblage is Pb–Zn–Cu–Ag (Au) and the sub-ore halo element assemblage is Bi–Co–Mo–Sn. The axial zoning sequence of the primary halo reveals an obvious reverse zoning feature and most of the reverse zone anomalies can be well explained by geology.

- The axial zoning indices, VECC, GP and the trilinear diagram of primary halo illustrate a probable hidden orebody when the supra- and sub-ore halos are superimposed. Fluctuations of the geochemical parameter reveal that the deposit has multiple-periodic and multiple staged overlapping characteristics. The F1 factor scores mainly reveal the later metamorphic-deformation and superimposed hydrothermal mineralization while the F2 factor scores represent exhalative-sedimentation mineralization.
- According to the axial primary halo and geology, four promising prospecting areas are delineated.

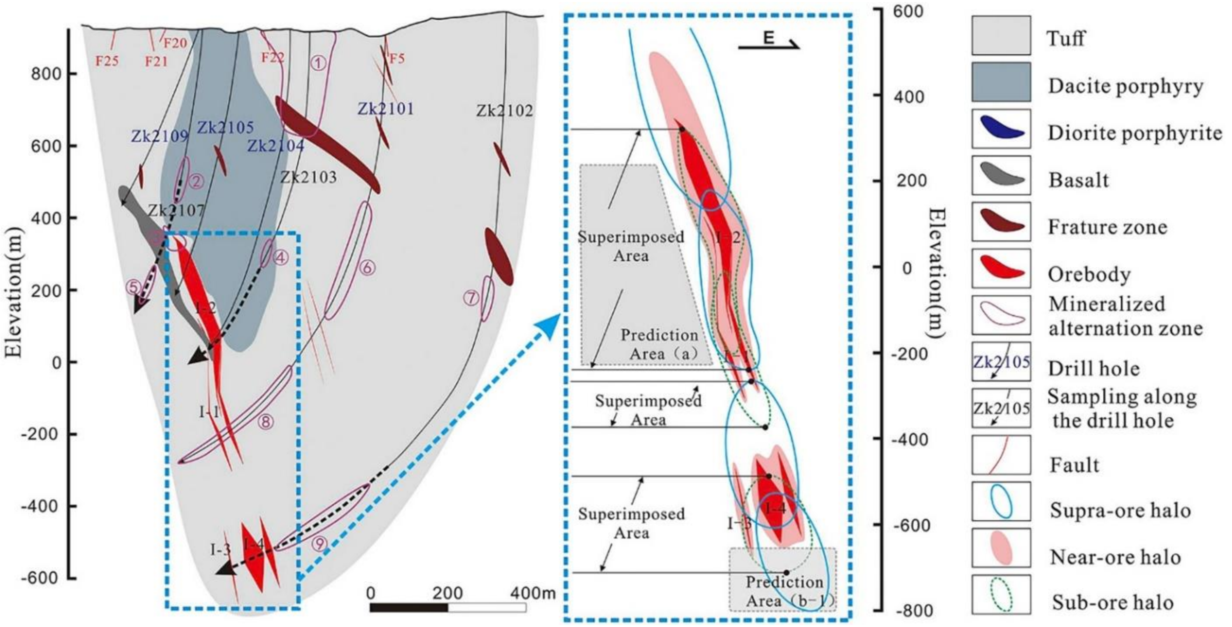


Fig. 2 Vertical superposition model map and deep prediction of primary halo in the No. 21 exploration line.

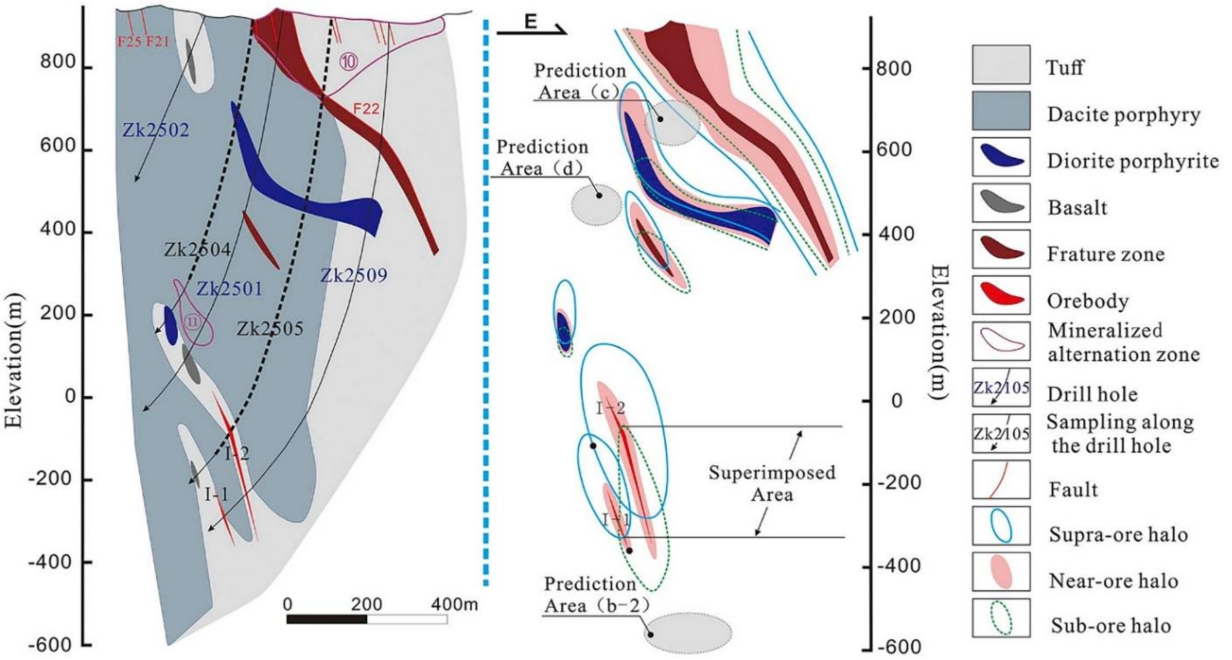


Fig. 3 Vertical superposition model map and deep prediction of primary halo in the No. 25 exploration line.

Contribution of the Ph.D. candidate

Partial data visualization and data interpretation

Section 6.2:
Gadolinium Contamination in Water Resources

Abstract

As a result of high doses of paramagnetic gadolinium (Gd) chelates administered in magnetic resonance imaging (MRI) exams, their unmetabolized excretion and insufficient removal in wastewater treatment plants (WWTPs), large amounts of anthropogenic Gd are released into surface water. The upward trend of gadolinium-based contrast agent (Gd-CA) administrations is expected to continue growing and consequently higher and higher anthropogenic Gd concentrations are annually recorded in water resources, which can pose a great threat to aquatic organisms and human beings. In addition, the feasibility of Gd retention in patients administered with Gd-CAs repeatedly and even potentially fatal diseases, including nephrogenic systemic fibrosis (NSF), due to trace amounts of Gd have recently arisen severe health concerns. Thus, there is a need to investigate probable adverse health effects of currently marketed Gd-CAs meticulously and to modify the actual approach in using Gd contrast media in daily practice in order to minimize unknown possible health risks. Furthermore, the employment of enhanced wastewater treatment processes that are capable of removing the stable contrast agents and the evaluation of the ecotoxicity of Gd chelates and human exposure to these emerging contaminants through dermal and ingestion pathways deserve more attention. On the other hand, point source releases of anthropogenic Gd into the aquatic environment presents the opportunity to assess surface water–groundwater interactions and trace the fate of wastewater plume as a proxy for the potential presence of other microcontaminants associated with treated wastewater in freshwater and marine systems.

1. Introduction

The upward trend of mining and using lanthanides (LAs: La to Lu) in the modern world has been accompanied by the release of similarly increasing amounts of those chemicals in the environment (Pedreira et al., 2018). Of the LAs, anthropogenic Gd (Gd_{anth}) mainly refers to the stable magnetic resonance imaging (MRI) Gd compounds that seems to be the most ubiquitous water contaminant worldwide. The Gd_{anth} inputs are mostly restricted to the regions with a highly developed healthcare system and mass application of MRI tests. Positive Gd anomalies of variable size have been observed in the influent and effluent of WWTPs, rivers, seawaters, groundwaters and tap waters (Lawrence and Bariel, 2010; Bau and Dulski, 1996; Hatje et al., 2016; Johannesson et al., 2017; Kulaksız and Bau, 2011a,b; Song et al., 2017; Nozaki et al., 2000; Zhu et al., 2004; Smith and Liu, 2018).

The LAs are either grouped into light LAs (LLAs), medium LAs (MLAs) and heavy LAs (HLAs) (Sholkovitz, 1995; Turetta et al., 2017), or only two subgroups of LLAs and HLAs (Song et al., 2017; Petrosino et al., 2013; Li et al., 2014). There is a difference between the organic complexation behavior of LLAs (La to Eu) and HLAs (Tb to Lu). At the junction of the two series of elements, Gd behaves as an HLA for the outer sphere (weak) complexation and as an LLA for the inner sphere (strong) complexation (Byrne and Li, 1995). The Gd_{anth} has provided opportunities for hydrological studies through tracing the treated wastewater plume (Lawrence and Bariel, 2010; Möller et al., 2002) while it has caused some community concerns because an increasing number of reports underlined the WWTP effluents as a source of micropollutants (Schwarzenbach et al., 2006; Ternes and Joss, 2007).

To date, no organized monitoring schedule for Gd complexes has resulted in a comprehensive understanding of their toxicity and regulation in the body and the environment. Hence, this review introduces anthropogenic Gd as an overlooked emerging microcontaminant that poses a threat to human and ecological health. The main objectives of the current investigation are to understand: (1) sources and fate of Gd_{anth} in the environment; (2) challenges versus opportunities of detecting Gd_{anth} in water resources; and (3) future research directions to fill the gaps in our understanding.

2. Natural versus anthropogenic Gd

The LA patterns of river waters with and without anthropogenic Gd, and wastewater samples with a remarkable positive Gd anomaly are visualized in Fig. 1. In the typical LA patterns for natural water, the relative concentrations of HLAs are higher than those of LLAs (Zhu et al., 2004). Moreover, there is a continuous trend from La to Gd prior to a decline to a lower level and another trend from Tb to Lu in the LA plot of pristine rivers such as Västerdaläven, Toshibetsu and Dhünn (Bau and Dulski, 1996) (Fig. 1a). The larger slopes from LLAs to HLAs might be due to the higher complexing abilities of HLAs with dissolved organic substances and the preferential incorporation of LLAs in the formation of secondary minerals (Zhu et al., 2004; Kulaksız and Bau, 2013; Haraguchi et al., 1998). However, in waters with less organic substances, submicron particles or colloids containing LLAs more than HLAs pass through the membrane filters with a pore size of 0.45 µm and lead to a flat LA distribution pattern (Haraguchi et al., 1998; Itoh et al., 2004).

In about 40% of MRI exams, 0.1 mmol of Gd per kg of body weight has been applied (Kanal, 2016; Clases et al., 2018) to improve clarity, contrast and differentiation of tissues in the MRI images. The figures for patients who undergo MRI exams (Fig. 2) and the global administration of Gd-CAs will continue to grow because of the increasing number of available MRI scanners worldwide. Therefore, a continuously increasing amount of Gd enters the aqueous environment and enriches Gd over LAs more and more (Fig. 1b and c) (Thomsen, 2017). In San Francisco Bay, increasing from ca. 10 pmol/kg in 1993 to 20 pmol/kg in 2000, anthropogenic Gd concentrations surged noticeably over the next 13 years and hit well above 100 pmol/kg in 2013. The proportion of anthropogenic Gd in seawater was surprisingly greater than 50% between 2001 and 2013 (Hatje et al., 2016). Generally, the inputs and concentrations of Gd_{anth} into the public sewage system and the environment vary greatly depending on the population size, level of access to health systems and the number of MRI imaging centers (Pedreira et al., 2018).

The acidification and filtration experiments suggest (Bau and Dulski, 1996; Knappe et al., 2005; Kulaksız and Bau, 2013): (1) a difference between the speciation of Gd_{anth} in contaminated waters and the speciation of geogenic LAs in non-contaminated waters; (2) presence of the excess Gd in the “dissolved” LA fraction (<0.2 µm) rather than the acid-soluble particulate fraction; and (3) negligible particle reactivity of the anomalous Gd in water resources. However, there is no guideline for regulating Gd-CAs in sewage effluents or receiving water bodies, partly because of the insufficient knowledge about the ecological effects of LAs (Pedreira et al., 2018).

3. Gd_{anth} and surface water–groundwater interactions

Since the early report of anthropogenic Gd anomalies in river waters (Bau and Dulski, 1996), similar positive anomalies have been reported in the dissolved load of many rivers worldwide (Smith and Liu, 2018; Song et al., 2017; de Campos and Enzweiler, 2016; Kulaksız and Bau, 2007, 2011a; Bau et al., 2006; Zhu et al., 2004; Nozaki et al., 2000; Möller et al., 2002). However, because Gd_{anth} is less particle-reactive and more conservative, the contaminant is expected to be negligibly removed by colloidal organic matter and suspended particles during the estuarine mixing of river water and seawater (Bau and Dulski, 1996; Nozaki et al., 2000; Censi et al., 2010). This further amplifies the positive Gd anomaly of the riverine LA input into seawater and negatively affects the LA signature in coastal waters and in sea basins with restricted water circulation (Bau and Dulski, 1996). The difference in magnitude observed between rivers and marine water systems is likely related to the hydrodynamics that efficiently transports contaminants and dilutes the Gd_{anth} signal in seawater (Pedreira et al., 2018).

On the other hand, the processes such as natural and induced bank filtration do not prevent migration of Gd_{anth} in the surface water resources into underground aquifers (Kulaksız and Bau, 2011a; Johannesson et al., 2017; Knappe et al., 2005; Möller et al., 2000; Rabiet et al., 2009). The changes in Gd_{anth} concentrations due to the mixing of bank filtrates and groundwater are well correlated with those of δD and $\delta^{18}O$, showing the suitability of Gd_{anth} for hydrological investigations at least on timescales of days to months (Möller et al., 2002). Studying groundwater hydrogeochemistry using stable Gd complexes with a high residence time in the environment is another opportunity (Bau and Dulski, 1996; Johannesson et al., 2017; Möller et al., 2000).

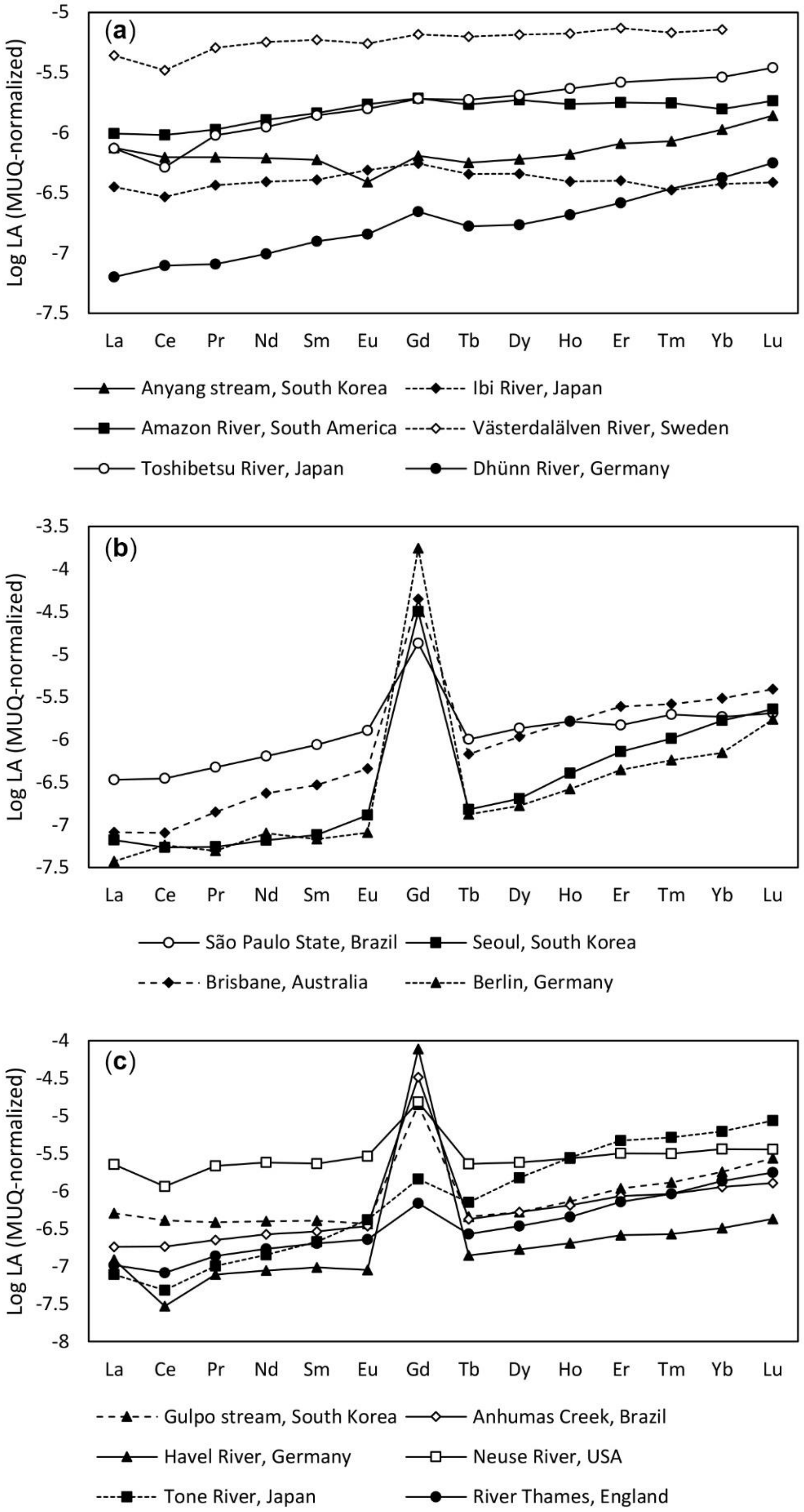


Fig. 1 MUQ-normalized LA patterns of: (a) river waters without Gd_{anth}; (b) WWTP effluents with pronounced positive Gd anomaly; (c) river waters with Gd_{anth}. The data are from Bau and Dulski (1996), Nozaki et al. (2000), Gaillardet et al. (2003), Zhu et al. (2004), Lawrence et al. (2009), Kulaksız and Bau (2011a), de Campos and Enzweiler (2016), Song et al. (2017) and Smith and Liu (2018).

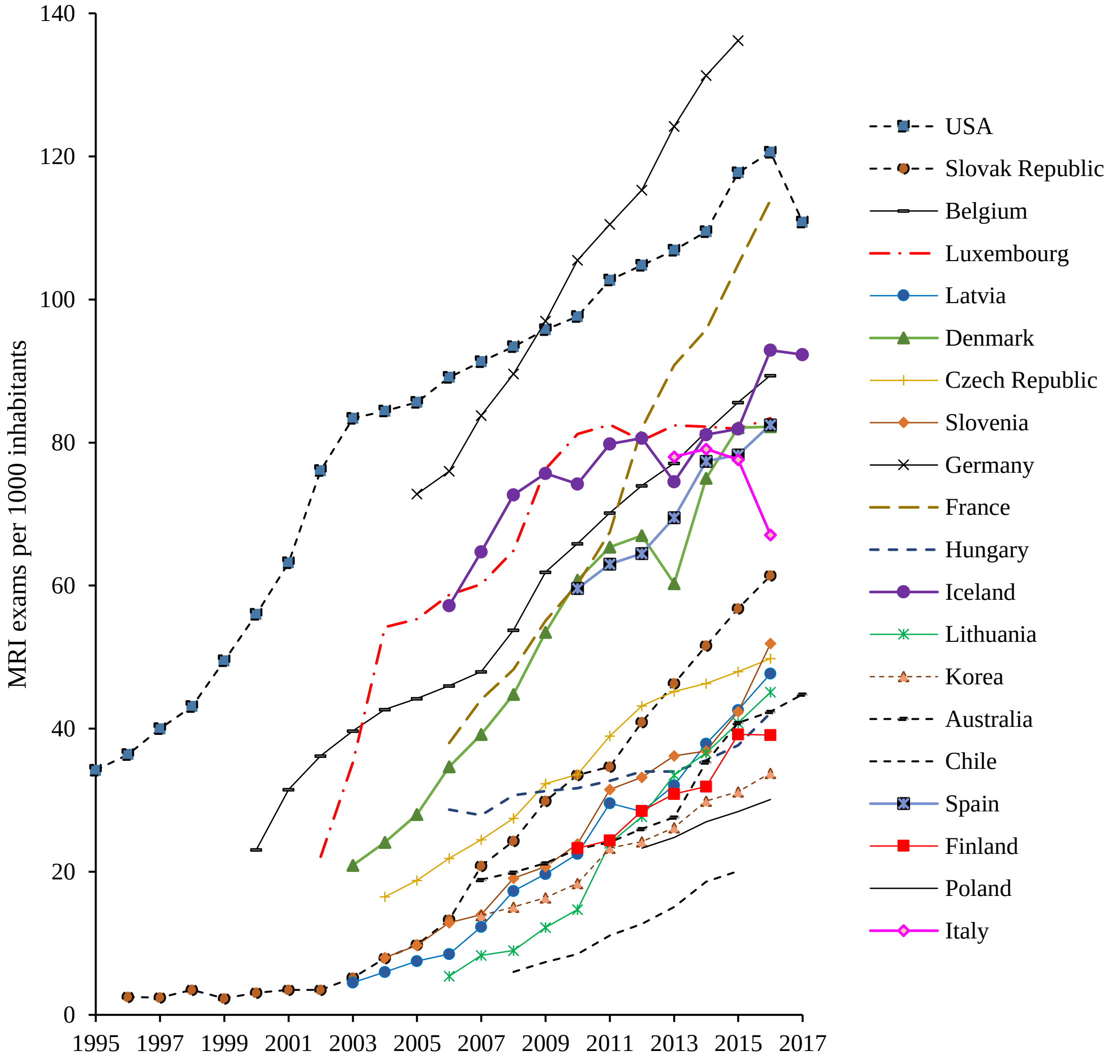


Fig. 2 Long-term trend of the number of MRI exams per 1000 inhabitants performed in a range of countries (OECD, 2018).

4. Health risks of Gd_{anth}

Gd accumulation in human tissues was recently uncovered following kidney failure (Ergün et al., 2006), nephrogenic systemic fibrosis (NSF) (Marckmann et al., 2007; Thakral et al., 2007), anaphylactic shock (Hasdenteufel et al., 2008) and eventually death (Kay, 2008). Patients administered with Gd-CAs may have incorporated part of the Gd chelates into the brain, bone and kidney, and even trace amounts of Gd can trigger potentially fatal diseases, including NSF (White et al., 2006; Darrah et al., 2009; Thakral and Abraham, 2009; US Food and Drug Administration, 2015; Rogosnitzky and Branch, 2016). On the other hand, high levels of Gd retention and NSF-like symptoms were experienced by some patients with normal renal function who underwent several MRI examinations (Sieber et al., 2008; Gathings et al., 2015; Semelka et al, 2016a, 2016b; Burke et al., 2016; Roberts et al., 2016). A considerable decrease of the number of NSF cases was observed after establishing a specific safety policy for the use of Gd-CAs by many scientific societies along with prominent health authorities (such as the Food and Drug Administration and the European Medicines Agency), and no new cases were reported after 2009, except for the rare isolated cases (Larson et al., 2015).

Recent studies have demonstrated Gd accumulation in the brain regions, liver, kidney, spleen, skin and bone of different animals (Kartamihardja et al., 2016; Wáng et al., 2015; Lohrke et al., 2017; McDonald et al., 2017; Bussi et al., 2018). Non-ionic linear Gd-CAs pose the highest risk for potential Gd retention (Birka et al., 2015; Sieber et al., 2008; Kartamihardja et al., 2016; Kanda et al., 2016; Murata et al., 2016). An investigation on the nonhuman primate placenta, fetal tissues and amniotic fluid (Oh et al., 2015) indicated that the Gd complex concentrations in the fetal tissues and amniotic fluid were minimal compared to the maternal injected dose, which may alleviate some concerns regarding MRI Gd chelate administration during pregnancy.

The mechanism of toxicity of lanthanides in plants depends on the La speciation, plant species and growth stage. LAs may interrupt the uptake of various essential nutrients and interfere with the corresponding functions (Gwenzi et al., 2018). The findings on Gd_{anth} in plants indicate the need for further investigation on the biological uptake and bioaccumulation of Gd-CAs by different plant species at different growth stages and consequently the biomagnification potential (Braun et al., 2018; Lindner et al., 2013; Lingott et al., 2016).

5. Knowledge gaps and future research directions

- Finding more effective techniques for wastewater treatment and water purification.
- Determining the degradation and transformation products of Gd-CAs during water treatment.
- Collecting an extensive spatiotemporal LA dataset to serve as a background level for monitoring purpose.
- Long-term bioaccumulation, stability and toxicity of Gd-CAs need to be addressed in the biosphere.
- Besides determination of total LA contents, various Gd-CAs should be considered to understand the environmental fate and potential health risk.
- Given that river water and seawater might be used for drinking purposes and recreational activities, human health risk assessment through chronic and acute exposure scenarios would be beneficial.
- Taking the opportunity to investigate water pollution and surface water–groundwater interactions.
- Further collaboration between medical scientists and environmental experts could be fruitful.

6. Conclusions

In this review, it was indicated that stable MRI contrast agents have been negligibly removed from wastewater and discharged into surface water, resulting in the detection of Gd_{anth} in river water, seawater, groundwater and tap water. Moreover, the number of MRI exams is on the increase worldwide, leading to greater Gd_{anth} concentrations in water resources in the future. Part of the administered Gd complexes have been incorporated in the brain, bone and kidney of patients and even trace amounts of them can encourage potentially fatal health problems such as NSF. Although paramagnetic Gd chelates have been detected as a microcontaminant in water resources since the mid-1990s, their uptake and bioaccumulation by different biota have not been thoroughly investigated and, as a result, no permissible limit has been introduced by authorities to screen safe food and water. It is probable that every individual takes up minute daily amounts of Gd via the consumption of contaminated drinking water and ingredients (meat, seafood and vegetables) that is considerably higher than the natural doses. Consequently, it is crucial to monitor Gd species in the environment regularly and promote life-long observations to assess their ecotoxicity and adverse human health effects. Cooperation between medical scientists and environmental experts is highly important to achieve further progress. The discrimination against bioaccumulation of nongeogenic Gd by at least some organisms and plants corroborates the previous evidence revealing long environmental half-lifes of the paramagnetic Gd chelates. Moreover, it adds to the growing body of evidence suggesting conservative anthropogenic Gd as a useful tracer of WWTP effluents in river water, lake water, seawater, groundwater and tap water. Evaluating the distribution of MRI Gd-CAs also offers a rather inexpensive yet robust way of monitoring xenobiotics such as pharmaceuticals and personal care products in water.

Contribution of the Ph.D. candidate

Determining research objectives; Data visualization; Data interpretation; Writing the original draft; Revising the draft

Section 6.3:

Bioavailability of Selected Major and Potentially Toxic Elements in Soil

Abstract

The pseudo-total concentrations and the bioavailable fractions of Al, Ca, Cu, K, Mg and Tl were determined in 1993 topsoil samples from Campania (southern Italy) by ICP-MS and ICP-ES techniques. The datasets were then compared through geochemical mapping and statistical analysis to highlight the spatial patterns and the key influencing factors. The medians for the bioavailable percentage of elements revealed the following decreasing order: Ca > K >> Mg ≃ Tl >> Cu >> Al. A multiple linear regression model was applied to the data, including grain size and organic matter content, and it was indicated that the pseudo-total concentration and grain size play a significant role in estimating the bioavailable fraction of most elements. The organic matter content is a significant input variable for prediction of the bioavailable Ca and K. In the future studies, it is recommended to consider more variables that govern bioavailability of elements for improving the estimations.

1. Introduction

The distribution and concentration of elements on the earth surface depend on the geological processes and the interference of anthropogenic activities with the natural processes. The elements of geogenic origin have generally strong bonds with the crystalline lattices of minerals while those of anthropogenic origin are more mobile. The latter is a key factor in soil remediation and the uptake of toxic metals by plants which facilitates their introduction in the food chain (Albanese, 2008). Under standard environmental conditions, the amount of a chemical that can be transferred from the environment to the biological context is defined as the bioavailable fraction of the total content (Petruzzelli et al., 2020; Ryu et al., 2010). Specifically, bioavailability is a process with three phases: (1) the "environmental availability" (also known as "external bioavailability") of an element which is mostly associated with its desorption from the matrix and accumulation in pore water; (2) the "environmental bioavailability" which is the fraction of element transferred from the matrix to the plant roots; and (3) the "toxicological bioavailability" which is related to accumulation of the element in plant tissues (Kim et al., 2015). In general, bioavailability depends on both soil characteristics (e.g., pH, clay percentage, organic matter, cation exchange capacity, redox potential and presence of oxides and hydroxides) and physiology of the life forms involved (National Research Council, 2002). Thus, the total (or pseudo-total) concentration of a potentially toxic element in soil should not be considered for assessing the adverse effects on biota and humans.

This study aims primarily at comparing the distribution patterns of the total and bioavailable concentrations of some major and potentially toxic elements in agricultural soils. The potential influence of some soil physicochemical properties, such as grain size and organic matter, on the spatial variability of the bioavailable concentrations will be evaluated as well.

2. Study area

Campania is an administrative region located along the Tyrrhenian margin of southern Italy. This territory is mainly made-up of volcanic lithotypes, carbonate rocks and sediments (Fig. 1a). The volcanism in the study area ranges from about 600 ka B.P. to the present (De Vivo et al., 2010; Peccerillo, 2020) and it is witnessed by the presence of four volcanic complexes: Mt. Roccamonfina, Mt Somma-Vesuvius, Phlegrean Fields and Ischia Island. Potassic to ultrapotassic rocks and pyroclastic deposits represent the main products of volcanic activity. Loose pyroclastic materials mantle large portions of the region and form cover beds on most carbonate reliefs and some siliciclastic formations (Fig. 1a).

The soil, particularly in the coastal plains, is characterized by a higher abundance of water and has an exceptional agricultural yield. Accordingly, Campania is one of the Italian regions where agriculture is flourishing and the territory is mostly occupied by herbaceous crops (e.g., wheat, cereals, weeds, pasture areas, ornamental plants and industrial crops), horticultural crops (e.g., potatoes and tomatoes) and orchards, combining chestnut groves, olive groves, vineyards, citrus groves and other minor fruit trees. Forests are mainly characterized by the presence of coniferous, broad-leaved, sclerophyllous, bushes and shrubs and they correspond with the Apennines relieves (Fig. 1b). The main urban areas include Naples metropolitan area, Benevento, Caserta and Avellino. The historical industrial areas of Campania are usually located close to the main cities and the recent ones are usually located in the easily accessible flat areas surrounded by territories with major agricultural use.

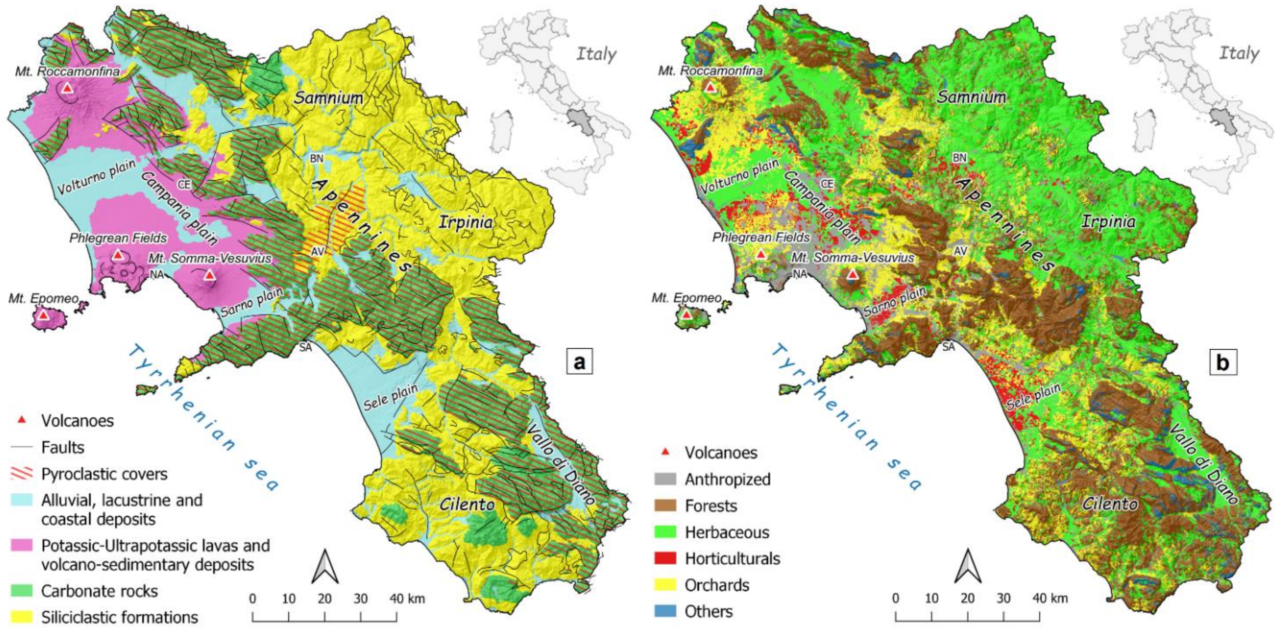


Fig. 1 (a) The simplified geological map and (b) the agricultural land use map.

3. Materials and methods

3.1. Sampling and analytical methods

Across the whole region, 7300 topsoil samples were collected from 0 to 15 cm below ground and concentrations of 53 elements were determined in the <2 mm fraction by inductively coupled plasma-mass spectrometry (ICP-MS) and inductively coupled plasma-emission spectrometry (ICP-ES) after aqua regia digestion (2 volumes of HCl + 2 volumes of HNO_3 + 2 parts of distilled water). Based on the analytical results, 1993 samples were then selected and the bioavailable fraction of elements was determined following ammonium nitrate (NH_4NO_3) treatment. Besides, the physicochemical properties were measured in 3316 samples (Nasta et al., 2020). The soil organic carbon content was determined using the dichromate method (Mebius, 1960) and converted into organic matter (OM). The mass fraction of the sand particles was also obtained by sieving whereas the clay and silt particles were measured by the hydrometer method (Gee and Or, 2002).

3.2. Data treatment, geochemical mapping and statistical analysis

For the aims of this study, only the samples underwent both ammonium nitrate and aqua regia leaching were considered. The elements of interest are Al, Ca, Cu, K, Mg and Tl because they are linked to agricultural practices (Horn, 1934; Prajapati and Modi, 2012; Rader et al., 2019; Rout et al., 2001; Sardans and Peñuelas, 2021; Thor, 2019; Yan and Hou, 2018; Yruela, 2005). The values below the detection limit were replaced by 50% of the detection limit (Reimann et al., 2008) and the Box-Cox transformation (Box and Cox, 1964) was implemented before imputation of the missing values using the Expectation-Maximization algorithm

(Dempster et al., 1977). The bioavailability percentage was calculated as a ratio between the bioavailable fraction (C_{AN}) and the pseudo-total concentration (C_{AR}) at each sample location.

The C_{AR}, C_{AN}, bioavailability percentage, OM and grain size were processed in a geographic information system environment and interpolated by Multifractal Inverse Distance Weighted (MIDW) method (Cheng, 1999). For separation of the pixels of the raster grids into discrete intervals, the concentration-area (C-A) method was subsequently applied (Albanese et al, 2007; Cheng et al., 1994; Zuo and Wang, 2020). Because there was not correspondence between the samples collected for C_{AN} and OM measurements, the OM values for the locations of the 1993 samples were extracted from the raster grid of OM to perform statistical analysis.

Multiple linear regression models were performed to understand the spatial features of the bioavailable concentration in soil. The C_{AN} value was set as the dependent (or response) variable whereas C_{AR}, OM and grain size were used as the independent (or explanatory) variables. The method refers to the standard stepwise regression, which consists of a series of iterations in which the explanatory variables are included in or excluded from the model step-by-step to achieve the optimal result (Reimann et al, 2008).

4. Results and discussion

4.1. Aluminum

The Al bioavailability percentage is generally very low with a median value of 0.0073 % (Fig. 2a) which represents a suitable condition regarding the agricultural use of most collected soil samples and the Al phytotoxicity. The low bioavailability percentage is probably due to the tendency of Al^{3+} to remain in the alteration environment as hydroxides and to participate in clay minerals. It is also noteworthy that the highest bioavailability percentages, up to 1.56%, are roughly related to the concomitant pyroclastic cover (Fig. 1a) and forest (Fig. 1b) where the soils are less alkaline compared to the rest of the territory (di Gennaro et al., 2002). The multiple linear regression model shows that Al_{AR} is a statistically significant predictor for Al_{AN}, but it explains only 5% of variability in the bioavailable aluminum (Table 1).

Table 1 The equations, adjusted R^2 and p-values of the multiple linear regression models. OM = organic matter; SN = sand; CL = clay; and SL = silt. The p-value range for the coefficients: *** = [0, 0.001]; ** = (0.001, 0.01]; * = (0.01, 0.05]; °°= (0.05, 0.1]; ° = (0.1, 1].

Regression equations	**R^2**	**p-values**
$Al_{AN} \simeq - 1.204e\text{-}04^{\circ\circ} + 1.534e\text{-}04^{***} Al_{AR}$	0.0468	<2.2e-16
$Ca_{AN} \simeq 1.575e\text{-}01^{***} + 4.858e\text{-}03^{***} Ca_{AR} + 8.943e\text{-}06^{**}$ OM + $5.572e\text{-}03^{***}$ CL	0.4126	<2.2e-16
$Cu_{AN} \simeq - 2.702e\text{-}01^{***} + 5.192e\text{-}03^{***} Cu_{AR} + 5.243e\text{-}03^{***}$ SN	0.7146	<2.2e-16
$K_{AN} \simeq 1.966e\text{-}01^{***} + 7.579e\text{-}03^{***} K_{AR} - 5.651e\text{-}06^{*}$ OM - $1.246e\text{-}03^{***}$ SN - $2.928e\text{-}03^{***}$ CL	0.1403	<2.2e-16
$Mg_{AN} \simeq 4.829e\text{-}02^{***} + 5.946e\text{-}03^{***} Mg_{AR} - 4.434e\text{-}04^{***}$ SN - $3.015e\text{-}04^{***}$ SL	0.2054	<2.2e-16
$Tl_{AN} \simeq - 9.14714e\text{-}02^{***} + 6.68643e\text{-}02^{***} Tl_{AR} + 1.7554e\text{-}03^{***}$ SL + $2.122e\text{-}04^{\circ}$ CL	0.3576	<2.2e-16

4.2. Calcium

The distribution patterns of Ca bioavailability percentage appear to be strongly different from those of pseudo-total and bioavailable concentrations. The highest values (up to 95%; Fig. 2b) occur in the southeastern sector of Mt. Roccamonfina, in Mt. Matese at the northern regional border and in Cilento. It is probable that, to some extent, the highest bioavailability percentage corresponds to the agricultural areas (e.g., southeastern sector of Mt. Roccamonfina) where soil pH and ion exchanges is regulated with some amendments (e.g., calcium oxide and hydroxide and raw product). The high Ca extractability likely reflects formation of soluble ions which are easily leachable. Ca_{AN}, OM and clay are the significant predictors for the bioavailable Ca fraction ($R^2 = 0.4$; Table 1) which agrees with the spatial distribution of the Ca_{AN}, clay and OM.

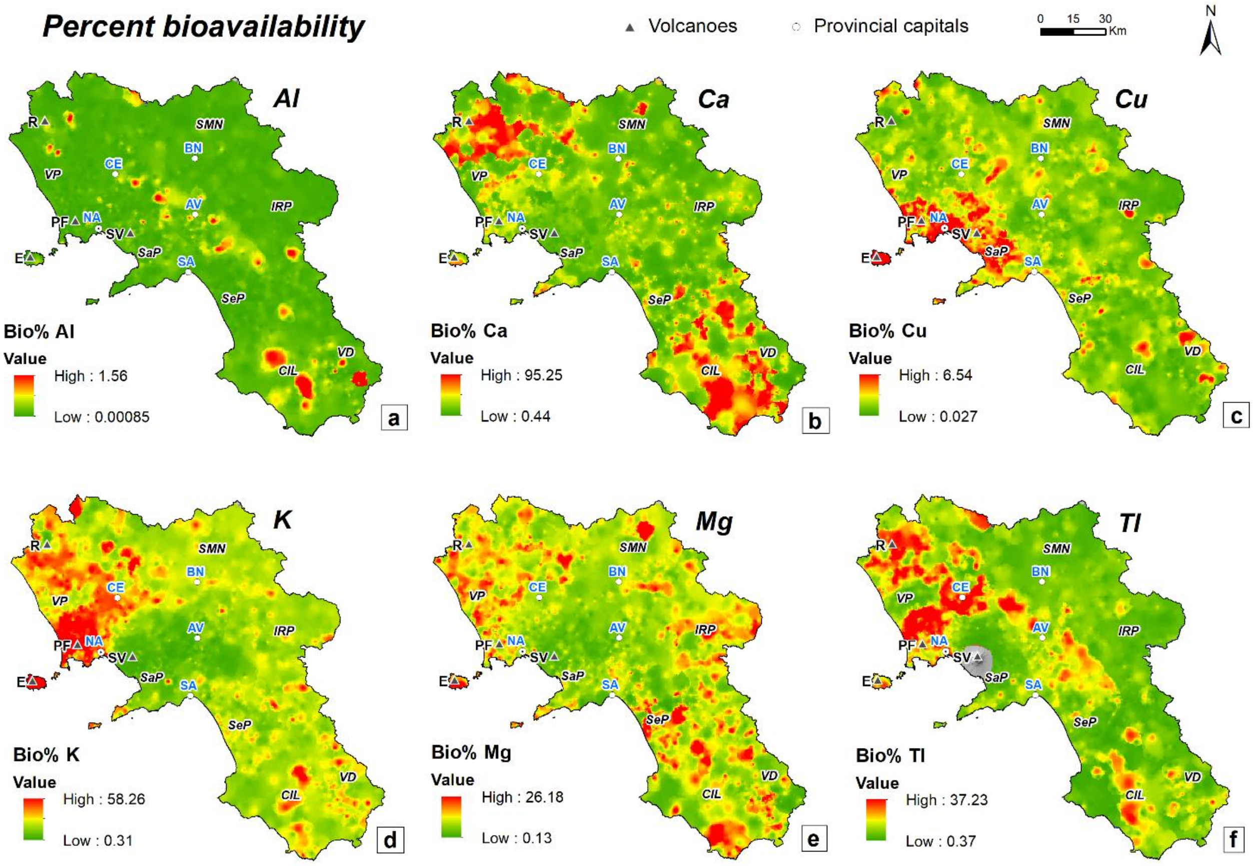


Fig. 2 The interpolated maps for the bioavailability percentage of (a) Al, (b) Ca, (c) Cu, (d) K, (e) Mg and (f) Tl. Volcanoes: R = Roccamonfina; PF = Phlegrean Fields; SV = Somma-Vesuvius; and E = Epomeo. River plains: VP = Volturno; SaP = Sarno; and SeP = Sele. Areas (in black): SMN = Samnium; IRP = Irpinia; CIL = Cilento; and VD = Vallo di Diano. Provincial capitals (in blue): NA = Napoli; CE = Caserta; AV = Avellino; BN = Benevento; and SA = Salerno.

4.3. Copper

Referring to the pseudo-total Cu, about 8% of samples overcome the national guideline value for agricultural use of soil (200 mg/kg; D.M. 46/2019). The Cu bioavailability percentage is generally low, but it shows marked values (up to 6.54%) in correspondence with the Vesuvian and Phlegrean territories (Fig. 2c) where the sand fraction is greater in soil, suggesting higher Cu extractability due to the low binding strength of the coarser fractions (Rieuwerts et al., 1998) and application of cupric fungicides in orchards (including vineyards) and horticulture in southern Campania plain (Fig. 1b). The multiple linear regression model (Table 1) indicates that Cu_{AR} and sand explain more than 70% of variability in Cu_{AN}.

4.4. Potassium

The K bioavailability percentage reaches the maximum values (up to 58%; Fig. 2d) in the volcanic areas of Phlegrean Fields, Mt. Roccamonfina and its surroundings, along with several small spots across the region. The potassium-based soil amendments are regular for cultivation in this sector of the study area. According to di Gennaro et al. (2002), the soils in the Phlegrean area and Mt. Roccamonfina are neutral and slightly acidic, respectively, which justify the greater K extractability. On the other hand, the sandy texture of the loose volcanic products surrounding Mt. Somma-Vesuvius probably leads to the lowest K bioavailability percentage because the element occurs in the crystalline structure of minerals. The overall multiple linear regression (Table 1) was found to be statistically significant although it has a low predictive ability ($R^2 = 0.14$). K_{AR}, organic matter, sand and clay are good predictors of K_{AN}. Excluding K_{AR} results in the negative β coefficients of the independent variables which means that there is an inverse proportionality with K bioavailability in soil.

4.5. Magnesium

The areas with the highest Mg bioavailability percentage are irregularly scattered across the study area (Fig. 2e), but they are mainly associated with the northern sector of Campania Plain and the siliciclastic formations in the inner and the southern sectors of the region. In general, there is a rough correspondence between the highest bioavailability percentage and the clayey soils. The Mg_{AR} values explain limited amounts of Mg_{AN} variability ($R^2 = 0.067$), but including sand and silt in the model enhances the R-squared value to 0.21. The negative signs of sand and clay coefficients imply that they have an inverse relationship with the changes in the bioavailable concentration of Mg.

4.6. Thallium

In more than 60% of the samples, the pseudo-total Tl concentration overcomes the national guideline values for agricultural soil (1 mg/kg; D.M. 46/2019). The spatial pattern of Tl bioavailability percentage (Fig. 2f) is similar to those of Tl_{AR} and Tl_{AN} geochemical maps and the areas reaching the highest values roughly correspond with the pyroclastic cover (Fig. 1a). The horticulture and orchard between the Volturno plain and the Phlegrean Fields also lead to high Tl bioavailability percentage. In general, Tl is considered mobile because it has a low ionic potential, the Fe oxyhydroxides do not easily adsorb Tl ions and it is hardly complexed with humic acids (Migaszewski and Gałuszka, 2021). The fitted multiple linear regression model (Table 1) is statistically significant although it is representative of limited variability ($R^2 = 0.36$). The Tl_{AR} and silt quantities account for the meaningful Tl_{AN} predictors.

5. Conclusions

A total of 1993 topsoil samples (0-15 cm depth) were collected in Campania region and concentrations of Al, Ca, Cu, K, Mg and Tl were detected in the <2 mm fraction by ICP-MS and ICP-ES after aqua regia digestion. The bioavailable fractions of elements were also determined following ammonium nitrate treatment. In another set of samples, the physicochemical properties were measured. The total/pseudo-total and bioavailable concentrations of elements generally highlight particular natural and/or anthropogenic features in the study area. The medians for the bioavailable percentage of elements follow the order of Ca > K >> Mg ≃ Tl >> Cu >> Al. Contrary to Al, Cu and Tl, the total/pseudo-total concentrations of Ca, K and Mg are poorly able to predict the bioavailable fraction. Including grain size and organic matter in the regression models substantially improved prediction of the bioavailable Ca, K, and Mg whereas they marginally explained bioavailability of Al, Cu, and Tl. The scientific literature shows that a wider range of factors could affect bioavailability and further research is required to consider more variables such as pH, redox potential together with Fe and Mn oxides in the model.

Contribution of the Ph.D. candidate

Partial data visualization and data interpretation

CHAPTER 7

Conclusions (Synthesis of the Key Results)

Conclusions (Synthesis of the Key Results)

The Phlegrean Fields, also known as Phlegraean Fields and Campi Flegrei, refer to a large active volcanic complex in south Italy with an eruptive history (mainly explosive) of at least 80 ka. The caldera is partially submerged in Pozzuoli Bay and shows evidence of bradysismic crisis. The structural setting of the volcano, eruptive history, precursors of volcanic activity and risk of volcanic eruption have been widely studied. The 40-ka CI and 15-ka NYT eruptions together with the associated collapses are generally reported as the main factors in forming the present structure of caldera. However, some researchers consider either NYT or CI as the caldera-forming eruption in Phlegrean Fields. Determining the caldera rims and studying hydrology/hydrogeochemistry of the local aquifer suffer from lack of a comprehensive insight into the volcanic structure. The environmental issues are of the utmost importance as well due to high population density and diversity of anthropogenic activities (urbanization, agricultural practices and industrial activity) in the study area.

The main research objective of this thesis is to characterize the dominant processes controlling geochemical composition of different environmental compartments in the study area. Briefly, the following points are dealt with in the thesis: (1) hydrogeochemistry of groundwater in the Phlegrean Fields volcanic aquifer; (2) the natural and anthropogenic factors regulating composition of the sediments from Pozzuoli Bay; and (3) topsoil geochemistry in Phlegrean Fields and the immediate surroundings. During the first phase of mobility restrictions due to COVID-19 pandemic, chemistry of Naples tap water was also investigated to gain some experience for data interpretation in the Ph.D. program.

Chapter 2 discusses groundwater composition in Phlegrean Fields. The first section is related to groundwater evolution, but the second one examines the dissolved gases along with major ions, lithium, arsenic and boron in underground water.

Section 2.1 presents concentrations of major ions, Li, As and B in 44 groundwater samples that are spatially distributed over the whole study area. Ratios of bicarbonate, chloride and sulfate are considered to perform hierarchical clustering and four groups are identified: (1) bicarbonate-rich groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The relationships between variables in each group is then explored using compositional data analysis and the classical methods. In comparison with the Piper diagram, the ilr-ion plot shows variations in major ions more clearly. Moreover, spurious correlation does not occur when the data are transformed from the simplex geometry to the Euclidean geometry. The scatterplot generated via the isometric log-ratios constructed by Na^{+}, Cl^{-} and HCO_3^{-} also reveals that the bicarbonate-rich meteoric water flows towards the sea, interacts with the chlorine-rich thermal/seawater endmember and gives rise to the mixed groundwater. The rising hydrothermal vapor, however, increases sulfate and arsenic quantities of the groundwater body around the Solfatara crater.

Section 2.2 is a follows-up study that explores radon and CO_2 quantities besides major ions, Li, As, B and physicochemical parameter (i.e. pH, TDS and temperature) in 26 groundwater samples to encourage filling the existing knowledge gap and provide more information for future investigations. Principal component analysis is applied to the clr-transfomed data and the results indicate that the groundwater is mainly derived from meteoric water and influenced by the CO_2-rich magmatic gas (in some cases) together with upflow of hydrothermal fluid and/or seawater through the fault and fracture system. Average concentration of radon in the chlorine-rich group is statistically higher than that of bicarbonate-rich group and spatial distribution of both radon and carbon dioxide highlights the areas with the highest contribution of the hydrothermal system. Dealing with constant sum constraint might provide new possibilities for monitoring volcanic activity through studying groundwater hydrogeochemistry.

Chapter 3 examines the geochemical fingerprint of Naples tap water to show water origin and the flow dynamics in a complex water distribution system. It could be interesting because quality of the local groundwater is not appropriate for human consumption and the tap water is supplied by the aquifers and well fields outside Phlegrean Fields. Both input water (i.e. the exploited groundwater for introduction in the distribution system) and tap water data obtained in 2017 were used and it was revealed that integrating compositional data analysis and multivariate statistics could be a valuable tool to seasonally monitor the prevalence of the different supply sources. It is worth mentioning that Naples distribution network has a historical stratified development and a large variety of materials were used for constructing pipelines and urban reservoirs, hindering setting up the advanced models for effective management. Therefore, the proposed approach may be used to preliminarily divide the water distribution system into unique clusters and encourage running water quality models depending on the data availability in each cluster. Performance of the technique could be improved by increasing the sampling density.

Chapter 4 explores the major processes that control sediment geochemistry in the eastern sector of Pozzuoli Bay where different sources have been proposed for the potentially toxic elements (particularly As) to date. In this chapter, the study area is divided into the proximal and distal zones before further evaluation. Meaningful surficial and vertical univariate patterns are observed for the clr-transformed values of metal(loid)s. The robust PCA characterized two strong associations in the proximal zone: (1) Hg, Cd, Cu, Pb, Zn, OM, mud, dist_coast and dist_seepage; and (2) Al, As, V, Fe, Cr, Ni, sand and dist_fumarole. The former is probably contributed by the industrial activity because the PTEs decrease along the cores and in the distal zone while the latter is likely attributed to the Phlegrean Fields pyroclastic deposits and the seepages with a hydrothermal component. Regarding arsenic geochemistry, it seems that submarine groundwater discharge, underwater fumaroles and the coastal landfill do not play a significant role. The geochemical association in distal zone is almost similar with proximal zone, but the low contribution of variables in robust PCA probably explains spatial occurrence of the main sources of elements in the proximal zone.

Chapter 5 evaluates PTEs and some other indicator elements for determining their probable origin in the <2-mm grains of 394 soil samples (0-20 cm depth). The correlation diagram and robust principal component analysis demonstrate that: (1) the Pb–Zn–Hg–Cd–Sb–Cr–Ni association probably highlights the anthropogenic origin (e.g., the heavy traffic load and fossil fuel combustion); (2) the Al–Fe–Mn–Ti–Tl–V–Co–As–U–Th association likely reveals the immobile elements during weathering of the pyroclastic deposits; and (3) the Na–K–B association probably indicates weathering of the potassium-rich volcanic products. For all age groups of children under 6, negligible non-carcinogenic health risk (certainty > 95%) is expected because of exposure to Pb and Zn in soil through ingestion. The CR_{ing} values for Pb are below 1E-06 and highlight negligible cancer risk as well. Nevertheless, exposure of children to lead through the inhalation pathway poses an acceptable cancer risk with the certainty of about 25% (in the 5-6 years old individuals) to 45% (in the 2-3 years age range). The certainty value is 90% for the children aged 0 to 1 year.

Chapter 6 refers to the collaborations with other researchers and it is composed of three sections. In the first one, applying multifractal analysis and statistical analysis for detecting elemental associations and orebody exploration was beneficial in the Ph.D. program (see the previous chapters). The second section shows that the gadolinium chelates injected for magnetic resonance imaging play an important role in contaminating water resources with anthropogenic Gd. More than 10 magnetic resonance imaging centers offer diagnostic imaging with contrast agents in Phlegrean Fields and the immediate surroundings (including Naples), but there is a lack of information on presence of anthropogenic Gd in the aqueous environment. In the last section, interpreting soil geochemical data in Campania region facilitated preparing the chapter on soil geochemistry in Phlegrean Fields. Joining all the activities improved the experience of working with other group members to achieve a research objective.

Section 6.1 investigates composition of 404 rock and ore samples for predicting deep Cu-Zn orebody in Ashele mining area (China). The element assemblages are As–Sb–Ba–Au for the supra-ore halo, Pb–Zn–Cu–Ag (Au) for the near-ore halo, and Bi–Co–Mo–Sn for the sub-ore halo. The axial zoning sequence of the primary halo reveals a reverse zoning feature, and most of the reverse zone anomalies can be well explained by the geology. Regarding the axial zoning indices, the vertical variation characteristics and the trilinear diagram of primary halo, a hidden orebody may lie under the known orebody where the supra-ore halo is superimposed on the sub-ore halo in the tail of the boreholes. Fluctuations of the geochemical parameter reveal the multiple-periodic and multiple-staged overlapping characters of the deposit. The F1 factor score map mainly shows the later metamorphic-deformation and superimposed hydrothermal mineralization, but the F2 factor score map represents exhalative-sedimentation mineralization. Finally, the vertical superposition model maps in Ashele Cu-Zn mining area were constructed and four promising prospecting areas were delineated regarding the metallogenic potential.

Section 6.2 explains that the stable magnetic resonance imaging contrast agents have been negligibly removed in wastewater treatment plants and resulted in contamination of river water, seawater, groundwater and tap water with anthropogenic Gd. Moreover, the number of diagnostic exams is on the increase across the globe, leading to introduction of higher anthropogenic Gd quantities in water resources in the future. Water microcontamination by the paramagnetic Gd chelates has been detected since the mid-1990s although uptake and bioaccumulation of Gd contrast agents by different biota have not been comprehensively investigated and no permissible limit has been proposed to screen safe food and water. Consequently, monitoring Gd species in the environment and promoting life-long observations to assess their ecotoxicity and human health effects are crucial. The collaboration of medical scientists and environmental experts significantly facilitates further progress. On the other hand, a growing body of evidence suggest the conservative anthropogenic Gd as a potential tracer of wastewater treatment plant effluents in river water, lake water, seawater, groundwater and tap water. Evaluating the spatial distribution of Gd chelates also offers a rather inexpensive yet robust way of monitoring xenobiotics in water.

Section 6.3 refers to geochemistry of 1993 topsoil samples (<15 cm depth) in Campania region. The pseudo-total concentrations and bioavailable fractions of Al, Ca, Cu, K, Mg and Tl were detected after aqua regia digestion and ammonium nitrate treatment, respectively. In another set of samples, the physicochemical properties were also measured. The spatial distribution of pseudo-total element concentrations and the bioavailable proportions is interpretable considering the natural and anthropogenic activities in the study area. The medians for the bioavailable percentage of elements follow the order of Ca > K >> Mg ≃ Tl >> Cu >> Al. Contrary to Al, Cu and Tl, the total/pseudo-total concentrations of Ca, K and Mg poorly predict the bioavailable fraction. Including grain size and organic matter in the linear regression models greatly improved prediction of the bioavailable Ca, K, and Mg, and slightly enhanced estimating bioavailability of Al, Cu, and Tl.

Based on the results of the thesis, the following research directions are proposed for further investigation: (1) implementing compositional data analysis for interpreting the hydrogeochemical data obtained during several decades of monitoring volcanic activity; (2) collecting more groundwater samples, determining a wider range of chemical variables (including Ra, U and Th), considering the principles of compositional data analysis and performing robust statistics for studying radon in the underground water; (3) setting up advanced hydraulic and water quality models for the pipe network in Naples to check if the map proposed in Chapter 3 is useful for engineers and managers; (4) performing high-resolution measurements along spatially dense sediment cores in east Pozzuoli Bay and applying the CoDA technique to reduce uncertainty in geochemical interpretations; (5) planning environmental monitoring and human biomonitoring studies in Naples to handle the public health concerns; (6) determining anthropogenic gadolinium concentration in groundwater and seawater to understand the possibility of organizing detailed research projects to shed light on some processes responsible for groundwater/seawater composition; and (7) taking a wider range of variables (such as pH, redox potential together with Fe and Mn oxides) into account for predicting the bioavailability of elements in soil.

REFERENCES

ABBaCo Project (2018). Sperimentazioni Pilota Finalizzate al Restauro Ambientale e Balneabilitá del SIN Bagnoli-Coroglio. Available online: http://www.szn.it/index.php/en/research/integrative-marine-ecology/research-projects-emi/abbaco. (Accessed on 28 October 2020).

AGIP (1987). Modello geotermico del sistemaflegreo. Sintesi, SEG-MESG Internal Report, 23 pp.

Aitchison, J. (1986). The Statistical Analysis of Compositional Data. Chapman & Hall, London.

Aitchison, J. (1983). Principal component analysis of compositional data. Biometrika, 70(1), 57-65. https://doi.org/10.1093/biomet/70.1.57.

Aitchison, J. (1982). The statistical analysis of compositional data. Journal of the Royal Statistical Society: Series B (Methodological), 44(2), 139-177. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x.

Aitchison, J., Barceló-Vidal, C., Egozcue, J. J., & Pawlowsky-Glahn, V. (2002). A concise guide for the algebraic-geometric structure of the simplex, the sample space for compositional data analysis. In Proceedings of IAMG (Vol. 2, pp. 387-392).

Aitchison, J., & Greenacre, M. (2002). Biplots of compositional data. Journal of the Royal Statistical Society: Series C (Applied Statistics), 51(4), 375-392. https://doi.org/10.1111/1467-9876.00275.

Aiuppa, A., Avino, R., Brusca, L., Caliro, S., Chiodini, G., D'Alessandro, W., ... & 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. https://doi.org/10.1016/j.chemgeo.2005.11.004.

Aiuppa, A., Fiorani, L., Santoro, S., Parracino, S., Nuvoli, M., Chiodini, G., ... & Tamburello, G. (2015). New ground-based lidar enables volcanic CO_2 flux measurements. Scientific Reports, 5(1), 1-12. https://doi.org/10.1038/srep13614.

Aiuppa, A., Moretti, R., Federico, C., Giudice, G., Gurrieri, S., Liuzzo, M., ... & Valenza, M. (2007). Forecasting Etna eruptions by real-time observation of volcanic gas composition. Geology, 35(12), 1115-1118. https://doi.org/10.1130/G24149A.1.

Aiuppa, A., Tamburello, G., Di Napoli, R., Cardellini, C., Chiodini, G., Giudice, G., ... & Pedone, M. (2013). First observations of the fumarolic gas output from a restless caldera: Implications for the current period of unrest (2005–2013) at Campi Flegrei. Geochemistry, Geophysics, Geosystems, 14(10), 4153-4169. https://doi.org/10.1002/ggge.20261.

Albanese, S. (2008). Evaluation of the bioavailability of potentially harmful elements in urban soils through ammonium acetate–EDTA extraction: a case study in southern Italy. Geochemistry: Exploration, Environment, Analysis, 8(1), 49-57. https://doi.org/10.1144/1467-7873/07-154.

Albanese, S., De Vivo, B., Lima, A., & Cicchella, D. (2007). Geochemical background and baseline values of toxic elements in stream sediments of Campania region (Italy). Journal of Geochemical Exploration, 93(1), 21-34. https://doi.org/10.1016/j.gexplo.2006.07.006.

Albanese, S., De Vivo, B., Lima, A., Cicchella, D., Civitillo, D., & Cosenza, A. (2010). Geochemical baselines and risk assessment of the Bagnoli brownfield site coastal sea sediments (Naples, Italy). Journal of Geochemical Exploration, 105(1-2), 19-33. https://doi.org/10.1016/j.gexplo.2010.01.007.

Albanese, S., Ebrahimi, P., Aruta, A., Cicchella, D., De Vivo, B., & Lima, A. (in preparation). Potentially toxic elements in the soils of Phlegrean Fields (south Italy) and the immediate surroundings: Spatial distribution, origin and probabilistic human health risk.

Albanese, S., Sadeghi, M., Lima, A., Cicchella, D., Dinelli, E., Valera, P., ... & Team, T. G. P. (2015). GEMAS: cobalt, Cr, Cu and Ni distribution in agricultural and grazing land soil of Europe. Journal of Geochemical Exploration, 154, 81-93. https://doi.org/10.1016/j.gexplo.2015.01.004.

Allocca, V., Coda, S., De Vita, P., Di Rienzo, B., Ferrara, L., Giarra, A., ... & 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. https://doi.org/10.1016/j.gexplo.2017.11.004.

Alonso, C., Castellote, M., Llorente, I., & Andrade, C. (2006). Ground water leaching resistance of high and ultra high performance concretes in relation to the testing convection regime. Cement and Concrete Research, 36(9), 1583-1594. https://doi.org/10.1016/j.cemconres.2006.04.004.

Álvarez-Vázquez, M. Á., Hošek, M., Elznicová, J., Pacina, J., Hron, K., Fačevicová, K., ... & Grygar, T. M. (2020). Separation of geochemical signals in fluvial sediments: New approaches to grain-size control and anthropogenic contamination. Applied Geochemistry, 123, 104791. https://doi.org/10.1016/j.apgeochem.2020.104791.

Aphalo, P.J. (2021). ggpmisc: Miscellaneous Extensions to 'ggplot2'. R package version 0.3.9. https://CRAN.R-project.org/package=ggpmisc.

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, 761. https://doi.org/10.1007/s10661-015-4988-2.

Arienzo, M., Donadio, C., Mangoni, O., Bolinesi, F., Stanislao, C., Trifuoggi, M., ... & Ferrara, L. (2017). Characterization and source apportionment of polycyclic aromatic hydrocarbons (pahs) in the sediments of gulf of Pozzuoli (Campania, Italy). Marine Pollution Bulletin, 124(1), 480-487. https://doi.org/10.1016/j.marpolbul.2017.07.006.

Arienzo, M., Toscanesi, M., Trifuoggi, M., Ferrara, L., Stanislao, C., Donadio, C., ... & Carella, F. (2019). Contaminants bioaccumulation and pathological assessment in Mytilus galloprovincialis in coastal waters facing the brownfield site of Bagnoli, Italy. Marine Pollution Bulletin, 140, 341-352. https://doi.org/10.1016/j.marpolbul.2019.01.064.

Armienti, P., Barberi, F., Bizojard, H., Clocchiatti, R., Innocenti, F., Metrich, N., ... & Sbrana, A. (1983). The Phlegraean Fields: magma evolution within a shallow chamber. Journal of Volcanology and Geothermal Research, 17(1-4), 289-311. https://doi.org/10.1016/0377-0273(83)90073-2.

Armiento, G., Caprioli, R., Cerbone, A., Chiavarini, S., Crovato, C., De Cassan, M., ... & Zaza, F. (2020). Current status of coastal sediments contamination in the former industrial area of Bagnoli-Coroglio (Naples, Italy). Chemistry and Ecology, 36(6), 579-597. https://doi.org/10.1080/02757540.2020.1747448.

Armienta, M. A., & De la Cruz-Reyna, S. (1995). Some hydro-geochemical fluctuations observed in Mexico related to volcanic activity. Applied Geochemistry, 10(2), 215-227. https://doi.org/10.1016/0883-2927(94)00044-7.

Aruta, A., Guarino, A., Ebrahimi, P., Dominech, S., Belyaeva, O., Tepanosyan, G., & Albanese, S. (2020). Radiological risk induced by road paving and building materials in the city of Naples, south Italy. In: Fiore, S. (Editor). I International Meeting of GeoHealth Scientists - GHC 2020, Bari, Italy. Digilabs Pub., Bari, Italy, pp. 51.

Avino, R., Capaldi, G., & Pece, R. (1999). Radon in active volcanic areas of Southern Italy. Il Nuovo Cimento C, 22(3-4), 379-385.

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. https://doi.org/10.1016/j.jvolgeores.2009.09.010.

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 Proceedings of the second United Nations symposium on the 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. https://doi.org/10.1016/0016-7037(88)90102-0.

Barberi, F., Corrado, G., Innocenti, F., & Luongo, G. (1984). Phlegraean Fields 1982–1984: brief chronicle of a volcano emergency in a densely populated area. Bulletin Volcanologique, 47(2), 175-185. https://doi.org/10.1007/BF01961547.

Barbieri, M., Franchini, S., Barberio, M. D., Billi, A., Boschetti, T., Giansante, L., ... & Stockmann, G. (2021). Changes in groundwater trace element concentrations before seismic and volcanic activities in Iceland during 2010–2018. Science of the Total Environment, 793, 148635. https://doi.org/10.1016/j.scitotenv.2021.148635.

Bartolín, H., Martínez, F., & Cortés, J. A. (2005). Topological GIS-based analysis of a water distribution network model. Applications of the minimum spanning tree. Computers and Control in the Water Industry (CCWI). https://doi.org/10.13140/RG.2.1.4996.2007.

Basu, K. (2016). Spatio temporal mapping of tuberculosis using log-empirical Bayesian kriging methods. Technical Report Indian Statistical Institute/Applied Statistics Unit.

Bau, M., & Dulski, P. (1996). Anthropogenic origin of positive gadolinium anomalies in river waters. Earth and Planetary Science Letters, 143(1-4), 245-255. https://doi.org/10.1016/0012-821X(96)00127-6.

Bau, M., Knappe, A., & Dulski, P. (2006). Anthropogenic gadolinium as a micropollutant in river waters in Pennsylvania and in Lake Erie, northeastern United States. Geochemistry, 66(2), 143-152. https://doi.org/10.1016/j.chemer.2006.01.002.

Beck, K. K., Mariani, M., Fletcher, M. S., Schneider, L., Aquino-López, M. A., Gadd, P. S., ... & Zawadzki, A. (2020). The impacts of intensive mining on terrestrial and aquatic ecosystems: A case of sediment pollution and calcium decline in cool temperate Tasmania, Australia. Environmental Pollution, 265, 114695. https://doi.org/10.1016/j.envpol.2020.114695.

Bergamin, L., Romano, E., Magno, M. C., Ausili, A., & Gabellini, M. (2005). Pollution monitoring of Bagnoli Bay (Tyrrhenian Sea, Naples, Italy), a sedimentological, chemical and ecological approach. Aquatic Ecosystem Health & Management, 8(3), 293-302. https://doi.org/10.1080/14634980500220866.

Beus, A.A., & Grigorian, S.V. (1977). Geochemical exploration methods for mineral deposits. Earth-Science Reviews, 14 (1), 67-69. https://doi.org/10.1016/0012-8252(78)90045-4.

Birka, M., Wentker, K. S., Lusmöller, E., Arheilger, B., Wehe, C. A., Sperling, M., ... & Karst, U. (2015). Diagnosis of nephrogenic systemic fibrosis by means of elemental bioimaging and speciation analysis. Analytical chemistry, 87(6), 3321-3328. https://doi.org/10.1021/ac504488k.

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(2), 281-295.

Bove, M. A., Ayuso, R. A., De Vivo, B., Lima, A., & Albanese, S. (2011). Geochemical and isotopic study of soils and waters from an Italian contaminated site: Agro Aversano (Campania). Journal of Geochemical Exploration, 109(1-3), 38-50. https://doi.org/10.1016/j.gexplo.2010.09.013.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211-243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x.

Braun, M., Zavanyi, G., Laczovics, A., Berényi, E., & Szabó, S. (2018). Can aquatic macrophytes be biofilters for gadolinium based contrasting agents?. Water Research, 135, 104-111. https://doi.org/10.1016/j.watres.2017.12.074.

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, 141, 1-5. https://doi.org/10.1016/j.gexplo.2014.03.022.

Buccianti, A., Mateu-Figueras, G., & Pawlowsky-Glahn, V. (Eds.). (2006). Compositional data analysis in the geosciences: from theory to practice. Geological Society of London. p. 224. https://doi.org/10.1144/GSL.SP.2006.264.

Burke, L. M., Ramalho, M., AlObaidy, M., Chang, E., Jay, M., & Semelka, R. C. (2016). Self-reported gadolinium toxicity: a survey of patients with chronic symptoms. Magnetic Resonance Imaging, 34(8), 1078-1080. https://doi.org/10.1016/j.mri.2016.05.005.

Bussi, S., Coppo, A., Botteron, C., Fraimbault, V., Fanizzi, A., De Laurentiis, E., ... & Maisano, F. (2018). Differences in gadolinium retention after repeated injections of macrocyclic MR contrast agents to rats. Journal of Magnetic Resonance Imaging, 47(3), 746-752. https://doi.org/10.1002/jmri.25822.

Byrne, R. H., & Li, B. (1995). Comparative complexation behavior of the rare earths. Geochimica et Cosmochimica Acta, 59(22), 4575-4589. https://doi.org/10.1016/0016-7037(95)00303-7.

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. https://doi.org/10.1016/j.gca.2007.04.007.

Canty, A., & Ripley, B. (2021). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-26.

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. https://doi.org/10.1007/BF00876715.

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. https://doi.org/10.1016/S0377-0273(96)00072-8.

Caracausi, A., Italiano, F., Paonita, A., Rizzo, A., & Nuccio, P. M. (2003). Evidence of deep magma degassing and ascent by geochemistry of peripheral gas emissions at Mount Etna (Italy): Assessment of the magmatic reservoir pressure. Journal of Geophysical Research: Solid Earth, 108(B10). https://doi.org/10.1029/2002JB002095.

Cardellini, C., Chiodini, G., Frondini, F., Avino, R., Bagnato, E., Caliro, S., ... & Rosiello, A. (2017). Monitoring diffuse volcanic degassing during volcanic unrests: the case of Campi Flegrei (Italy). Scientific Reports, 7(1), 1-15. https://doi.org/10.1038/s41598-017-06941-2.

Carroll, J. J., Mather, A. E. (1991). IUPAC-NIST Solubility Database. Department of Chemical Engineering, University of Alberta, Edmonton, Alberta, Canada. https://srdata.nist.gov/solubility/sol_detail.aspx?goBack=Y&sysID=62_1.

Celico, P., De Gennaro, M., Pagano, D., Ronca, A., Stanzione, D., & Vallario, A. (1987). In: 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, pp. 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.

Celico, P., Esposito, L., de Gennaro, M., & Mastrangelo, E. (1994). La falda ad Oriente della città di Napoli: Idrodinamica e qualità delle acque. Geologica Romana, 30, 653-660.

Censi, P., Zuddas, P., Randazzo, L. A., Saiano, F., Mazzola, S., Arico, P., ... & Punturo, R. (2010). Influence of dissolved organic matter on rare earth elements and yttrium distributions in coastal waters. Chemistry and Ecology, 26(2), 123-135. https://doi.org/10.1080/02757541003627720.

Chandrasekharam, D., & Bundschuh, J. (2002). Geochemistry of thermal waters and thermal gases. Geothermal Energy Resources for Developing Countries, 253-267.

Chayes, F. (1960). On correlation between variables of constant sum. Journal of Geophysical research, 65(12), 4185-4193.

Chen, Y. (1996). Ore-forming Conditions and Metallogenic Prognosis of the Ashele Copper-Zinc Metallogenic Belt, Xinjiang, China, 1st ed. Geological Publishing House, Beijing (330 pp).

Chen, Y., & Zhao, P. (1998). Zonation in primary halos and geochemical prospecting pattern for the Guilaizhuang gold deposit, eastern China. Nonrenewable Resources, 7(1), 37-44. https://doi.org/10.1007/BF02782507.

Cheng, Q. (1999). Spatial and scaling modelling for geochemical anomaly separation. Journal of Geochemical exploration, 65(3), 175-194. https://doi.org/10.1016/S0375-6742(99)00028-X.

Cheng, Q., Agterberg, F. P., & Ballantyne, S. B. (1994). The separation of geochemical anomalies from background by fractal methods. Journal of Geochemical Exploration, 51(2), 109-130. https://doi.org/10.1016/0375-6742(94)90013-2.

Chiodini, G., Frondini, F., Cardellini, C., Granieri, D., Marini, L., & Ventura, G. (2001). CO_2 degassing and energy release at Solfatara volcano, Campi Flegrei, Italy. Journal of Geophysical Research: Solid Earth, 106(B8), 16213-16221. https://doi.org/10.1029/2001JB000246.

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). https://doi.org/10.1029/2002GL016790.

Chiodini, G., Caliro, S., Aiuppa, A., Avino, R., Granieri, D., Moretti, R., & Parello, F. (2011). First $^{13}C/^{12}C$ isotopic characterisation of volcanic plume CO_2. Bulletin of volcanology, 73(5), 531-542. https://doi.org/10.1007/s00445-010-0423-2.

Chiodini, G., Caliro, S., Avino, R., Bini, G., Giudicepietro, F., De Cesare, W., ... & Tripaldi, S. (2021). Hydrothermal pressure-temperature control on CO_2 emissions and seismicity at Campi Flegrei (Italy). Journal of Volcanology and Geothermal Research, 414, 107245. https://doi.org/10.1016/j.jvolgeores.2021.107245.

Chiodini, G., Caliro, S., Cardellini, C., Granieri, D., Avino, R., Baldini, A., ... & 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). https://doi.org/10.1029/2008JB006258.

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. https://doi.org/10.1130/G33251.1.

Chiodini, G., Vandemeulebrouck, J., Caliro, S., D'Auria, L., De Martino, P., Mangiacapra, A., & Petrillo, Z. (2015). Evidence of thermal-driven processes triggering the 2005–2014 unrest at Campi Flegrei caldera. Earth and Planetary Science Letters, 414, 58-67. https://doi.org/10.1016/j.epsl.2015.01.012.

Chirkov, A. M. (1975). Radon as a possible criterion for predicting eruptions as observed at Karymsky volcano. Bulletin Volcanologique, 39(1), 126-131. https://doi.org/10.1007/BF02596952.

Cicchella, D., De Vivo, B., & Lima, A. (2005). Background and baseline concentration values of elements harmful to human health in the volcanic soils of the metropolitan and provincial areas of Napoli (Italy). Geochemistry: Exploration, Environment, Analysis, 5(1), 29-40. https://doi.org/10.1144/1467-7873/03-042.

Cicchella, D., Giaccio, L., Dinelli, E., Albanese, S., Lima, A., Zuzolo, D., ... & De Vivo, B. (2015). GEMAS: Spatial distribution of chemical elements in agricultural and grazing land soil of Italy. Journal of Geochemical Exploration, 154, 129-142. https://doi.org/10.1016/j.gexplo.2014.11.009.

Ciotoli, G., Voltaggio, M., Tuccimei, P., Soligo, M., Pasculli, A., Beaubien, S. E., & Bigi, S. (2017). Geographically weighted regression and geostatistical techniques to construct the geogenic radon potential map of the Lazio region: A methodological proposal for the European Atlas of Natural Radiation. Journal of Environmental Radioactivity, 166, 355-375. https://doi.org/10.1016/j.jenvrad.2016.05.010.

Civita, M. (1969). Idrogeologia del massiccio del Terminio-Tuoro (Campania), Memorie e note dell'Istituto di Geologia applicate dell'Università di Napoli.

Clases, D., Sperling, M., & Karst, U. (2018). Analysis of metal-based contrast agents in medicine and the environment. TrAC Trends in Analytical Chemistry, 104, 135-147. https://doi.org/10.1016/j.trac.2017.12.011.

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. https://doi.org/10.1016/0022-1694(78)90044-6.

D'Auria, L., Giudicepietro, F., Martini, M., & Lanari, R. (2012). The 4D imaging of the source of ground deformation at Campi Flegrei caldera (southern Italy). Journal of Geophysical Research: Solid Earth, 117(B8), 1-13. https://doi.org/10.1029/2012JB009181.

D’Auria, L., Pepe, S., Castaldo, R., Giudicepietro, F., Macedonio, G., Ricciolino, P., ... & 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. https://doi.org/10.1038/srep13100.

Daliri, M., Somma, R., Buccino, M., Troise, C., Molisso, F., & De Natale, G. (2020). Numerical Analysis of Arsenic Concentration in the Marine Sediments of Bagnoli Bay. The 30th International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers. 2020, September.

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.

Damiani, V., Baudo, R., De Rosa, S., De Simone, R., Ferretti, O., Izzo, G., & Serena, F. (1987). A case study: Bay of Pozzuoli (gulf of Naples, Italy). In: Thomas, R.L., Evans, R., Hamilton, A.L., Munawar, M., Reynoldson, T.B., Sadar, M.H. (Eds.), Ecological Effects of in Situ Sediment Contaminants, Developments in Hydrobiology, vol. 39. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-4053-6_21.

Darrah, T. H., Prutsman-Pfeiffer, J. J., Poreda, R. J., Ellen Campbell, M., Hauschka, P. V., & Hannigan, R. E. (2009). Incorporation of excess gadolinium into human bone from medical contrast agents. Metallomics, 1(6), 479-488. https://doi.org/10.1039/b905145g.

de Campos, F. F., & Enzweiler, J. (2016). Anthropogenic gadolinium anomalies and rare earth elements in the water of Atibaia River and Anhumas Creek, Southeast Brazil. Environmental Monitoring and Assessment, 188(5), 281. https://doi.org/10.1007/s10661-016-5282-7.

de Campos, F. F., & Licht, O. A. B. (2021). Correlation diagrams: Graphical visualization of geochemical associations using the EzCorrGraph app. Journal of Geochemical Exploration, 220, 106657.

De Moor, J. M., Aiuppa, A., Pacheco, J., Avard, G., Kern, C., Liuzzo, M., ... & Fischer, T. P. (2016). Short-period volcanic gas precursors to phreatic eruptions: Insights from Poás Volcano, Costa Rica. Earth and Planetary Science Letters, 442, 218-227. https://doi.org/10.1016/j.epsl.2016.02.056.

De Natale, G., Troise, C., Mark, D., Mormone, A., Piochi, M., Di Vito, M. A., ... & Somma, R. (2016). The Campi Flegrei Deep Drilling Project (CFDDP): New insight on caldera structure, evolution and hazard implications for the Naples area (Southern Italy). Geochemistry, Geophysics, Geosystems, 17(12), 4836-4847. https://doi.org/10.1002/2015GC006183.

De Natale, G., Troise, C., Piochi, M., Mormone, A., Aiello, G., Barra, D., Di Vito, M.A., Isaia, R., Mark, D.F., Smith, C.V., Wiersberg, T., Somma, R., Erzinger, J., & Jurado Rodriguez, M.J. (2014). The CFDDP 506 m drilling activity at Campi Flegrei (Italy): volcanological and environmental clues on the eastern sector of the caldera from analyses of mud log and cores. 1st International Workshop on Volcano Geology (Madeira, Portugal).

De Pippo, T., Donadio, C., Pennetta, M., Terlizzi, F., Vecchione, C., & Vegliante, M. (2002). Seabed morphology and pollution along the Bagnoli coast (Naples, Italy): a hypothesis for environmental restoration. Marine Ecology, 23, 154-168. https://doi.org/10.1111/j.1439-0485.2002.tb00015.x.

De Pippo, T., Pescatore, T., & Vecchione, C. (1988). Caratteri granulometrici dei sedimenti dei terrazzi del Golfo di Pozzuoli. Memorie della Societa Geologica Italiana, 41, 1005-1014.

De Siena, L., Del Pezzo, E., & Bianco, F. (2010). Seismic attenuation imaging of Campi Flegrei: Evidence of gas reservoirs, hydrothermal basins, and feeding systems. Journal of Geophysical Research: Solid Earth, 115(B9), 1-18. https://doi.org/10.1029/2009JB006938.

De Vita, P., Allocca, V., Celico, F., Fabbrocino, S., Mattia, C., Monacelli, G., Musilli, … & Celico, P. (2018). Hydrogeology of continental southern Italy. Journal of Maps, 14(2), 230-241. https://doi.org/10.1080/17445647.2018.1454352.

De Vivo, B., Belkin, H., & Lima, A. (Eds.). (2008). Environmental Geochemistry: Site Characterization, Data Analysis and Case Histories. Elsevier, Oxford.

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 Netherlands, Amsterdam, pp. 356-384.

De Vivo, B., Lima, A., Albanese, S. & Cicchella, D. (2006). Atlante geochimico-ambientale della Regione Campania. Aracne Editrice, Roma.

De Vivo, B., Lima, A., Albanese, S., Cicchella, D., Rezza, C., Civitillo, D., Minolfi, G., & Zuzolo, D. (2016). Atlante geochimico-ambientale dei suoli della Campania (Environmental Geochemical Atlas of Campania Soils). Aracne Editrice, Roma, 978-88- 548-9744-1pp. 364 (in Italian).

De Vivo, B., Petrosino, P., Lima, A., Rolandi, G., & Belkin, H. E. (2010). Research progress in volcanology in the Neapolitan area, southern Italy: a review and some alternative views. Mineralogy and Petrology, 99(1), 1-28. https://doi.org/10.1007/s00710-009-0098-6.

De Vivo, B., Rolandi, G., Gans, P. B., Calvert, A., Bohrson, W. A., Spera, F. J., & Belkin, H. E. (2001). New constraints on the pyroclastic eruptive history of the Campanian volcanic Plain (Italy). Mineralogy and Petrology, 73(1), 47-65. https://doi.org/10.1007/s007100170010.

Deino, A. L., Orsi, G., de Vita, S., & Piochi, M. (2004). The age of the Neapolitan Yellow Tuff caldera-forming eruption (Campi Flegrei caldera–Italy) assessed by 40Ar/39Ar dating method. Journal of Volcanology and Geothermal Research, 133(1-4), 157-170. https://doi.org/10.1016/S0377-0273(03)00396-2.

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. https://doi.org/10.1016/j.jvolgeores.2010.05.014.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39(1), 1-22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

Deuerlein, J., Elhay, S., & Simpson, A. R. (2016). Fast graph matrix partitioning algorithm for solving the water distribution system equations. Journal of Water Resources Planning and Management, 142(1), 04015037. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000561.

Deuerlein, J. W. (2008). Decomposition model of a general water supply network graph. Journal of Hydraulic Engineering, 134(6), 822-832. https://doi.org/10.1061/(ASCE)0733-9429(2008)134:6(822).

di Gennaro, A., Aronne, G., De Mascellis, R., Vingiani, S., Sarnataro, M., Abalsamo, P., Cona, F., Vitelli, L., & Arpaia, G. (2002). I sistemi di terre della Campania. Monografia e carta 1:250.000, con legenda. SELCA, Firenze, 63 pp.

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. https://doi.org/10.1007/BF02597225.

Di Napoli, R., Aiuppa, A., Sulli, A., Caliro, S., Chiodini, G., Acocella, V., ... & Valenza, M. (2016). Hydrothermal fluid venting in the offshore sector of Campi Flegrei caldera: A geochemical, geophysical, and volcanological study. Geochemistry, Geophysics, Geosystems, 17(10), 4153-4178.

Di Nardo, A., & Di Natale, M. (2010). A design support methodology for district metering of water supply networks. Proceedings of the 12th Annual Conference on Water Distribution Systems Analysis (WDSA), American Society of Civil Engineers, Tucson, AZ, pp. 870–887. https://doi.org/10.1061/41203(425)80.

Di Vito, M. A., Isaia, R., Orsi, G., Southon, J. D., De Vita, S., d'Antonio, M., ... & Piochi, M. (1999). Volcanism and deformation since 12,000 years at the Campi Flegrei caldera (Italy). Journal of Volcanology and Geothermal Research, 91(2-4), 221-246. https://doi.org/10.1016/S0377-0273(99)00037-2.

Di Vito, M., Lirer, L., Mastrolorenzo, G., & Rolandi, G. (1987). The 1538 Monte Nuovo eruption (Campi Flegrei, Italy). Bulletin of Volcanology, 49(4), 608-615. https://doi.org/10.1007/BF01079966.

Diao, K., Farmani, R., Fu, G., Astaraie-Imani, M., Ward, S., & Butler, D. (2014). Clustering analysis of water distribution systems: identifying critical components and community impacts. Water Science and Technology, 70(11), 1764-1773. https://doi.org/10.2166/wst.2014.268.

Diao, K., Zhou, Y., & Rauch, W. (2013). Automated creation of district metered area boundaries in water distribution systems. Journal of Water Resources Planning and Management, 139(2), 184-190. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000247.

D.Lgs.152/2006. Decreto Legislativo 3 aprile 2006, n. 152 “Norme in materia ambientale”. Gazzetta Ufficiale n. 88 del 14-4-2006, Supplemento Ordinario n.96. http://www.camera.it/parlam/leggi/deleghe/06152dl.htm.

D.M. 99/97, Regolamento sui criteri e sul metodo in base ai quali valutare le perdite degli acquedotti e delle fognature, Gazzetta Ufficiale, 1997, n. 90.

Dmitrijeva, M., Ehrig, K. J., Ciobanu, C. L., Cook, N. J., Verdugo-Ihl, M. R., & Metcalfe, A. V. (2019). Defining IOCG signatures through compositional data analysis: A case study of lithogeochemical zoning from the Olympic Dam deposit, South Australia. Ore Geology Reviews, 105, 86-101. https://doi.org/10.1016/j.oregeorev.2018.12.013.

Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S., & Albanese, S. (2020a). Potentially Toxic Elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (Southern Italy): a focus on cadmium. Proceedings of I International Meeting of GeoHealth Scientists 35-40. https://doi.org/10.14644/ghc2020.006.

Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S., & Albanese, S. (2020b). Potentially toxic elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (southern Italy). In: Fiore, S. (Ed.), I International Meeting of GeoHealth Scientists GHC 2020 BARI. Italy. Digilabs Pub., Bari, Italy, p. 51.

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. https://doi.org/10.7185/gold2021.6064.

Dominech, S., Yang, S., Aruta, A., Gramazio, A., & Albanese, S. (2022). Multivariate analysis of dilution-corrected residuals to improve the interpretation of geochemical anomalies and determine their potential sources: The Mingardo River case study (Southern Italy). Journal of Geochemical Exploration, 232, 106890. https://doi.org/10.1016/j.gexplo.2021.106890.

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. https://doi.org/10.1016/j.gexplo.2014.11.008.

Ducci, D., & Tranfaglia, G. (2008). Effects of climate change on groundwater resources in Campania (southern Italy). Geological Society, London, Special Publications, 288(1), 25-38. https://doi.org/10.1144/SP288.3.

Duggal, V., Sharma, S., & Mehra, R. (2017). Radon levels in drinking water of Fatehabad district of Haryana, India. Applied Radiation and Isotopes, 123, 36-40.

Dvorak, J. J., & Mastrolorenzo, G. (1991). The mechanisms of recent vertical crustal movements in Campi Flegrei caldera, southern Italy. Geological Society of America, Vol. 263, Colorado, USA.

Ebrahimi, P. (2013). Environmental Geochemistry and Bioavailability of Heavy Metals in Coastal and Marine Sediments of Chabahar Bay, SE Iran. Master's thesis. Shiraz University, Shiraz, Iran. https://doi.org/10.13140/RG.2.2.17519.66722.

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. https://doi.org/10.1039/D0EW01123A.

Ebrahimi, P., & Barbieri, M. (2019). Gadolinium as an emerging microcontaminant in water resources: threats and opportunities. Geosciences, 9(2), 93. https://doi.org/10.3390/geosciences9020093.

Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., ... & Albanese, S. (2022). Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy). Journal of Geochemical Exploration, 106922. https://doi.org/10.1016/j.gexplo.2021.106922.

Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Cicchella, D., & Albanese, S. (under review). The relationship between dissolved gases (radon and carbon dioxide) and other geochemical signatures in Phlegraean Fields volcanic aquifer: a follow-up study.

Egozcue, J. J., & Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. Mathematical Geology, 37(7), 795-828. https://doi.org/10.1007/s11004-005-7381-9.

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. https://doi.org/10.1023/A:1023818214614.

Ergün, I., Keven, K., Uruc, I., Ekmekci, Y., Canbakan, B., Erden, I., & Karatan, O. (2006). The safety of gadolinium in patients with stage 3 and 4 renal failure. Nephrology Dialysis Transplantation, 21(3), 697-700. https://doi.org/10.1093/ndt/gfi304.

ESRI (2016). ArcGIS Help Library: What is Empirical Bayesian Kriging? In: Environmental Systems Research Institute, I. (Ed.), ArcGIS Help Library. ESRI, United States of America.

EU Directive 98/83/EC, The quality of water intended for human consumption, Official Journal of the European Communities: Legis., 1998, 3(L 330), 32-54.

Farmani, R., & Butler, D. (2014). Implications of urban form on water distribution systems performance. Water Resources Management, 28(1), 83-97. https://doi.org/10.1007/s11269-013-0472-3.

Faucon, P., Adenot, F., Jacquinot, J. F., Petit, J. C., Cabrillac, R., & Jorda, M. (1998). Long-term behaviour of cement pastes used for nuclear waste disposal: review of physico-chemical mechanisms of water degradation. Cement and Concrete Research, 28(6), 847-857. https://doi.org/10.1016/S0008-8846(98)00053-2.

Faucon, P., Adenot, F., Jorda, M., & Cabrillac, R. (1997). Behaviour of crystallised phases of Portland cement upon water attack. Materials and Structures, 30(8), 480-485. https://doi.org/10.1007/BF02524776.

Field, A. (2013). Discovering statistics using IBM SPSS statistics. SAGE, London.

Filzmoser, P. (1999). Robust principal component and factor analysis in the geostatistical treatment of environmental data. Environmetrics: The Official Journal of the International Environmetrics Society, 10(4), 363-375. https://doi.org/10.1002/(SICI)1099-095X(199907/08)10:4%3C363::AID-ENV362%3E3.0.CO;2-0.

Filzmoser, P., Hron, K., & Reimann, C. (2009). Principal component analysis for compositional data with outliers. Environmetrics: The Official Journal of the International Environmetrics Society, 20(6), 621-632. https://doi.org/10.1002/env.966.

Filzmoser, P., Hron, K., & Templ, M. (2018). Applied Compositional Data Analysis: With Worked Examples in R. Springer Nature, Switzerland. https://doi.org/10.1007/978-3-319-96422-5.

Fisher, R. V., Orsi, G., Ort, M., & Heiken, G. (1993). Mobility of a large-volume pyroclastic flow – emplacement of the Campanian ignimbrite, Italy. Journal of Volcanology and Geothermal Research, 56(3), 205-220. https://doi.org/10.1016/0377-0273(93)90017-L.

Flood, R. P., Bloemsma, M. R., Weltje, G. J., Barr, I. D., O'Rourke, S. M., Turner, J. N., & Orford, J. D. (2016). Compositional data analysis of Holocene sediments from the West Bengal Sundarbans, India: Geochemical proxies for grain-size variability in a delta environment. Applied Geochemistry, 75, 222-235. https://doi.org/10.1016/j.apgeochem.2016.06.006.

Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. Biometrika, 58(3), 453-467. https://doi.org/10.1093/biomet/58.3.453.

Gaillardet, J., Viers, J., & Dupré, B. (2003). Trace elements in river waters. In Holland, H. D., & Turekian, K. K. Treatise on geochemistry, Elsevier, pp. 225-272. https://doi.org/10.1016/B0-08-043751-6/05165-3.

Gao, Z., Fang, W., Hu, R., & Li, W. (2010). Tectonic setting and geochemical characteristics of dacite andesite-basalt volcanic rocks in Ashele copper deposit, Xinjiang. Mineral Deposits, 29(2), 218-229.

Gathings, R. M., Reddy, R., Santa Cruz, D., & Brodell, R. T. (2015). Gadolinium-associated plaques: a new, distinctive clinical entity. JAMA Dermatology, 151(3), 316-319. https://doi.org/10.1001/jamadermatol.2014.2660.

Gauthier, P. J., Condomines, M., & Hammouda, T. (1999). An experimental investigation of radon diffusion in an anhydrous andesitic melt at atmospheric pressure: Implications for radon degassing from erupting magmas. Geochimica et Cosmochimica Acta, 63(5), 645-656. https://doi.org/10.1016/S0016-7037(98)00305-6.

Gee, G.W., & Or, D. (2002). Particle-size analysis. In: Dane, J.H., & Topp, G.C., Eds., Methods of Soil Analysis, Part 4, Physical Methods. SSSA, Book Series 5, Madison, pp. 255-293.

Gérard, B. (2000). Hydrolysis of cement-based materials: A review. Pore Solution in Hardened Cement Paste. Aedificatio Publishers, Freiburg, 271-304.

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. (Ed.), Application of Geochemistry in Geothermal Reservoir Development. UNITAR, pp. 119-144.

Giglioli, S., Colombo, L., Contestabile, P., Musco, L., Armiento, G., Somma, R., ... & Azzellino, A. (2020). Source apportionment assessment of marine sediment contamination in a post-industrial area (Bagnoli, Naples). Water, 12(8), 2181. https://doi.org/10.3390/w12082181.

Giudicepietro, F., Chiodini, G., Caliro, S., De Cesare, W., Esposito, A. M., Galluzzo, D., ... & Vandemeulebrouck, J. (2019). Insight into Campi Flegrei caldera unrest through seismic tremor measurements at Pisciarelli fumarolic field. Geochemistry, Geophysics, Geosystems, 20(11), 5544-5555. https://doi.org/10.1029/2019GC008610.

Giugni, M., Fontana, N., Portolano D., & Romanelli, D. (2008). A DMA Design for "Napoli Est" Water Distribution System, Proceedings of the 13th World Water Congress.

Giustolisi, O., & Ridolfi, L. (2014). New modularity-based approach to segmentation of water distribution networks. Journal of Hydraulic Engineering, 140(10), 04014049. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000916.

Gong, Q., Yan, T., Li, J., Zhang, M., & Liu, N. (2016). Experimental simulation of element mass transfer and primary halo zone on water-rock interaction. Applied Geochemistry, 69, 1-11. https://doi.org/10.1016/j.apgeochem.2016.04.001.

Gong, X., & Richman, M. B. (1995). On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. Journal of Climate, 8(4), 897-931. https://www.jstor.org/stable/26199911.

Granieri, D., Chiodini, G., Avino, R., & Caliro, S. (2014). Carbon dioxide emission and heat release estimation for Pantelleria Island (Sicily, Italy). Journal of Volcanology and Geothermal Research, 275, 22-33. https://doi.org/10.1016/j.jvolgeores.2014.02.011.

Graziano, R. S., Gozzi, C., & Buccianti, A. (2020). Is Compositional Data Analysis (CoDA) a theory able to discover complex dynamics in aqueous geochemical systems?. Journal of Geochemical Exploration, 211, 106465. https://doi.org/10.1016/j.gexplo.2020.106465.

Grezzi, G., Ayuso, R. A., De Vivo, B., Lima, A., & Albanese, S. (2011). Lead isotopes in soils and groundwaters as tracers of the impact of human activities on the surface environment: The Domizio-Flegreo Littoral (Italy) case study. Journal of Geochemical Exploration, 109(1-3), 51-58. https://doi.org/10.1016/j.gexplo.2010.09.012.

Groves-Kirkby, C. J., Denman, A. R., Crockett, R. G. M., Phillips, P. S., & Gillmore, G. K. (2006). Identification of tidal and climatic influences within domestic radon time-series from Northamptonshire, UK. Science of the Total Environment, 367(1), 191-202. https://doi.org/10.1016/j.scitotenv.2005.11.019.

Guagliardi, I., Zuzolo, D., Albanese, S., Lima, A., Cerino, P., Pizzolante, A., ... & Cicchella, D. (2020). Uranium, thorium and potassium insights on Campania region (Italy) soils: Sources patterns based on compositional data analysis and fractal model. Journal of Geochemical Exploration, 212, 106508. https://doi.org/10.1016/j.gexplo.2020.106508.

Guarino, A., Albanese, S., Cicchella, D., Ebrahimi, P., Dominech, S., Allocca, C., Romano, N., De Vivo, B., & Lima, A. (submitted). Selected major and potentially toxic elements bioavailability in agricultural soils of Campania region (Italy): The spatial patterns and influencing factors.

Guarino, A., Cicchella, D., Lima, A., & Albanese, S. (2022). Radon flux estimates, from both gamma radiation and geochemical data, to determine sources, migration pathways, and related health risk: The Campania region (Italy) case study. Chemosphere, 287, 132233. https://doi.org/10.1016/j.chemosphere.2021.132233.

Guglielminetti, M. (1986). Mofete geothermal field. Geothermics, 15(5-6), 781-790.

Güler, C., Thyne, G. D., McCray, J. E., & Turner, K. A. (2002). Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeology Journal, 10(4), 455-474. https://doi.org/10.1007/s10040-002-0196-6.

Gupta, S. K., Vollmer, M. K., & Krebs, R. (1996). The importance of mobile, mobilisable and pseudo total heavy metal fractions in soil for three-level risk assessment and risk management. Science of the Total Environment, 178(1-3), 11-20.

Gwenzi, W., Mangori, L., Danha, C., Chaukura, N., Dunjana, N., & Sanganyado, E. (2018). Sources, behaviour, and environmental and human health risks of high-technology rare earth elements as emerging contaminants. Science of the Total Environment, 636, 299-313. https://doi.org/10.1016/j.scitotenv.2018.04.235.

Haddad, S., Restieri, C., & Krishnan, K. (2001). Characterization of age-related changes in body weight and organ weights from birth to adolescence in humans. Journal of Toxicology and Environmental Health Part A, 64(6), 453-464. https://doi.org/10.1080/152873901753215911.

Haraguchi, H., Itoh, A., Kimata, C., & Miwa, H. (1998). Speciation of yttrium and lanthanides in natural water by inductively coupled plasma mass spectrometry after preconcentration by ultrafiltration and with a chelating resin. Analyst, 123(5), 773-778. https://doi.org/10.1039/A708253C.

Hasdenteufel, F., Luyasu, S., Renaudin, J. M., Paquay, J. L., Carbutti, G., Beaudouin, E., ... & Kanny, G. (2008). Anaphylactic shock after first exposure to gadoterate meglumine: two case reports documented by positive allergy assessment. Journal of Allergy and Clinical Immunology, 121(2), 527-528. https://doi.org/10.1016/j.jaci.2007.08.027.

Hatje, V., Bruland, K. W., & Flegal, A. R. (2016). Increases in anthropogenic gadolinium anomalies and rare earth element concentrations in San Francisco Bay over a 20 year record. Environmental Science & Technology, 50(8), 4159-4168. https://doi.org/10.1021/acs.est.5b04322.

Helba, H. A., Khalil, K. I., Mo'nes, M. M., & Khalek, I. A. A. (2020). Zonation in primary geochemical haloes for orogenic vein-type gold mineralization in the Quartz Ridge prospect, Sukari gold mine area, Eastern Desert of Egypt. Journal of Geochemical Exploration, 209, 106378. https://doi.org/10.1016/j.gexplo.2019.106378.

Herrera, M., Canu, S., Karatzoglou, A., Pérez-García R., & Izquierdo, J. (2010). An approach to water supply clusters by semisupervised learning. International Environmental Modelling and Software Society (iEMSs), Ottawa.

Hidalgo, A., Andrade C., & Alonso, C., (2000). An accelerated leaching test to evaluate the long term behaviour of concrete in waste disposal. 5th CANMET/ACI International Conference on Durability of Concrete, Barcelona, p. 373-381.

Horn, E. E., (1934). The effect of thallium on plant growth. Circular N.409, United States Department of Agriculture.

Hosseini-Dinani, H., & Aftabi, A. (2016). Vertical lithogeochemical halos and zoning vectors at Goushfil Zn–Pb deposit, Irankuh district, southwestern Isfahan, Iran: Implications for concealed ore exploration and genetic models. Ore Geology Reviews, 72, 1004-1021. https://doi.org/10.1016/j.oregeorev.2015.09.023.

Huang, J., Wu, Y., Sun, J., Li, X., Geng, X., Zhao, M., ... & Fan, Z. (2021). Health risk assessment of heavy metal (loid) s in park soils of the largest megacity in China by using Monte Carlo simulation coupled with Positive matrix factorization model. Journal of Hazardous Materials, 415, 125629. https://doi.org/10.1016/j.jhazmat.2021.125629.

İçhedef, M., Giammanco, S., Neri, M., Catalano, R., Immè, G., Morelli, D., ... & Giudice, N. (2020). In soil radon anomalies and volcanic activity on Mt. Etna (Italy). Journal of Environmental Radioactivity, 218, 106267. https://doi.org/10.1016/j.jenvrad.2020.106267.

ICRAM (2005). Caratterizzazione ambientale dei fondali e degli arenili inclusi nella perimetrazione del sito di bonifica di Napoli Bagnoli-Coroglio. Relazione tecnica conclusiva. BoI-Pr-CA-BA-relazione-02.04. Il Commissario di Governo per le Bonifiche e Tutela delle Acque nella regione Campania delegato ex O.P.C.M. 2425/96 e successive, p. 113.

Imperato, M., Adamo, P., Naimo, D., Arienzo, M., Stanzione, D., & Violante, P. (2003). Spatial distribution of heavy metals in urban soils of Naples city (Italy). Environmental Pollution, 124(2), 247-256. https://doi.org/10.1016/S0269-7491(02)00478-5.

Ippolito, F., Ortolani, F., & Russo, M. (1973). Struttura marginale dell'Appennino campano: reinterpretazioni di dati di antiche ricerche di idrocarburi. Memorie della Societa Geologica Italiana. 12, 127–250.

Isaia, R., Di Vito, M. A., De Vita, S., Rosi, M., & Sbrana, A. (2012). Comment on "40 Ar/39 Ar dating of tuff vents in the Campi Flegrei caldera (southern Italy): toward a new chronostratigraphic reconstruction of the Holocene volcanic activity" by Fedele et al.[Bull Volcanol; 73: 1323-1336]. Bulletin of Volcanology, 74(1), 293-296. https://doi.org/10.1007/s00445-011-0561-1.

ISPRA (Istituto superiore per la protezione e la ricerca ambientale) (2018). Carta Geologica d'Italia. 1:50,000 Sheets: 446-447 Napoli.

ISTAT (2019). Tavole mortalitá: Speranza di vita alla nascita. Available online: https://dati.istat.it/index.aspx?queryid=7283# (Accessed on 31 January 2022).

Italian Legislative Decree 31/2001, Attuazione della direttiva 98/83/CE relativa alla qualità delle acque destinate al consumo umano, Gazzetta Ufficiale, 2001, vol. 52.

Itoh, A., Nagasawa, T., Zhu, Y., Lee, K. H., Fujimori, E., & Haraguchi, H. (2004). Distributions of major-to-ultratrace elements among the particulate and dissolved fractions in natural water as studied by ICP-AES and ICP-MS after sequential fractionation. Analytical Sciences, 20(1), 29-36. https://doi.org/10.2116/analsci.20.29.

Izquierdo, J., Herrera, M., Montalvo I., & Pérez-García, R. (2009). Agent-based division of water distribution systems into distric metered areas. Proceedings of the 4th International Conference on Software and Data Technologies, SciTePress - Science and and Technology Publications, pp. 83-90. https://doi.org/10.5220/0002237800830090.

Johannesson, K. H., Palmore, C. D., Fackrell, J., Prouty, N. G., Swarzenski, P. W., Chevis, D. A., ... & Burdige, D. J. (2017). Rare earth element behavior during groundwater–seawater mixing along the Kona Coast of Hawaii. Geochimica et Cosmochimica Acta, 198, 229-258. https://doi.org/10.1016/j.gca.2016.11.009.

Jordan, G., Petrik, A., De Vivo, B., Albanese, S., Demetriades, A., Sadeghi, M., & Team, T. G. P. (2018). GEMAS: Spatial analysis of the Ni distribution on a continental-scale using digital image processing techniques on European agricultural soil data. Journal of Geochemical Exploration, 186, 143-157. https://doi.org/10.1016/j.gexplo.2017.11.011.

Kabata-Pendias, A., & Mukherjee, A. B. (2007). Trace elements from soil to human. Springer. Berlin.

Kamarehie, B., Jafari, A., Zarei, A., Fakhri, Y., Ghaderpoori, M., & Alinejad, A. (2019). Non-carcinogenic health risk assessment of nitrate in bottled drinking waters sold in Iranian markets: a Monte Carlo simulation. Accreditation and Quality Assurance, 24(6), 417-426. https://doi.org/10.1007/s00769-019-01397-5.

Kanal, E. (2016). Gadolinium based contrast agents (GBCA): Safety overview after 3 decades of clinical experience. Magnetic Resonance Imaging, 34(10), 1341-1345. https://doi.org/10.1016/j.mri.2016.08.017.

Kanda, T., Nakai, Y., Oba, H., Toyoda, K., Kitajima, K., & Furui, S. (2016). Gadolinium deposition in the brain. Magnetic Resonance Imaging, 34(10), 1346-1350. https://doi.org/10.1016/j.mri.2016.08.024.

Kartamihardja, A. A. P., Nakajima, T., Kameo, S., Koyama, H., & Tsushima, Y. (2016). Impact of impaired renal function on gadolinium retention after administration of gadolinium-based contrast agents in a mouse model. Investigative Radiology, 51(10), 655-660. https://doi.org/10.1097/RLI.0000000000000295.

Kay, J. (2008). Nephrogenic systemic fibrosis: a gadolinium-associated fibrosing disorder in patients with renal dysfunction. Annals of the Rheumatic Diseases, 67, iii66-iii69. http://dx.doi.org/10.1136/ard.2008.102475.

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.

Kim, R. Y., Yoon, J. K., Kim, T. S., Yang, J. E., Owens, G., & Kim, K. R. (2015). Bioavailability of heavy metals in soils: definitions and practical implementation—a critical review. Environmental Geochemistry and Health, 37(6), 1041-1061. https://doi.org/10.1007/s10653-015-9695-y.

Knappe, A., Möller, P., Dulski, P., & Pekdeger, A. (2005). Positive gadolinium anomaly in surface water and ground water of the urban area Berlin, Germany. Geochemistry, 65(2), 167-189. https://doi.org/10.1016/j.chemer.2004.08.004.

Koike, K., Yoshinaga, T., Suetsugu, K., Kashiwaya, K., & Asaue, H. (2015). Controls on radon emission from granite as evidenced by compression testing to failure. Geophysical Journal International, 203(1), 428-436. https://doi.org/10.1093/gji/ggv290.

Kresic, N., & Mikszewski, A. (2012). Hydrogeological conceptual site models: data analysis and visualization. CRC press, Boca Raton, FL.

Krivoruchko, K., & Gribov, A. (2014). Pragmatic Bayesian Kriging for Non-Stationary and Moderately Non-Gaussian Data. In: Pardo-Igúzquiza, E. et al. (Eds.), Mathematics of Planet Earth: Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 61-64. https://doi.org/10.1007/978-3-642-32408-6_15.

Krivoruchko, K. (2012). Empirical Bayesian Kriging: implemented in ArcGIS Geoestatistical Analyst. ArcUser Fall 2012.

Krivoruchko, K., & Gribov, A. (2019). Evaluation of empirical Bayesian kriging. Spatial Statistics, 32, 100368.

Krumbein, W. C., & and Graybill, F. (1965). An introduction to statistical models in geology; McGraw-Hill Book Co., New York, 475 pp.

Kulaksız, S., & Bau, M. (2013). Anthropogenic dissolved and colloid/nanoparticle-bound samarium, lanthanum and gadolinium in the Rhine River and the impending destruction of the natural rare earth element distribution in rivers. Earth and Planetary Science Letters, 362, 43-50. https://doi.org/10.1016/j.epsl.2012.11.033.

Kulaksız, S., & Bau, M. (2011a). Anthropogenic gadolinium as a microcontaminant in tap water used as drinking water in urban areas and megacities. Applied Geochemistry, 26(11), 1877-1885. https://doi.org/10.1016/j.apgeochem.2011.06.011.

Kulaksız, S., & Bau, M. (2011b). Rare earth elements in the Rhine River, Germany: first case of anthropogenic lanthanum as a dissolved microcontaminant in the hydrosphere. Environment International, 37(5), 973-979. https://doi.org/10.1016/j.envint.2011.02.018.

Kulaksız, S., & Bau, M. (2007). Contrasting behaviour of anthropogenic gadolinium and natural rare earth elements in estuaries and the gadolinium input into the North Sea. Earth and Planetary Science Letters, 260(1-2), 361-371. https://doi.org/10.1016/j.epsl.2007.06.016.

Kumar, A., Walia, V., Yang, T. F., Fu, C. C., Singh, S., Bajwa, B. S., & Arora, V. (2016). Soil ^{222}Rn concentration, CO_2 and CH_4 flux measurements around the Jwalamukhi area of North-West Himalayas, India. Radiation Protection Dosimetry, 171(2), 262-266. https://doi.org/10.1093/rpd/ncw071.

Kynčlová, P., Hron, K., & Filzmoser, P. (2017). Correlation between compositional parts based on symmetric balances. Mathematical Geosciences, 49(6), 777-796. https://doi.org/10.1007/s11004-016-9669-3.

Lagerblad, B. (1999). Texture and chemistry of historic concrete subjected to prolonged hydration. Workshop on water in cement paste and concrete hydration and pore structure, Nordic Concrete Federation, Skagen, p. 148-154.

Landauer, R. S. (1961). The Radioactive Disintegration Constant, Lambda. Radiology, 77(2), 306-306. https://doi.org/10.1148/77.2.306b.

Larson, K. N., Gagnon, A. L., Darling, M. D., Patterson, J. W., & Cropley, T. G. (2015). Nephrogenic systemic fibrosis manifesting a decade after exposure to gadolinium. JAMA Dermatology, 151(10), 1117-1120. https://doi.org/10.1001/jamadermatol.2015.0976.

Lawrence, M. G., & Bariel, D. G. (2010). Tracing treated wastewater in an inland catchment using anthropogenic gadolinium. Chemosphere, 80(7), 794-799. https://doi.org/10.1016/j.chemosphere.2010.05.001.

Lawrence, M. G., Ort, C., & Keller, J. (2009). Detection of anthropogenic gadolinium in treated wastewater in South East Queensland, Australia. Water Research, 43(14), 3534-3540. https://doi.org/10.1016/j.watres.2009.04.033.

Li, H. (1998). Primary halo model for buried ore prospecting of large and super-large gold deposits. Metallurgical Industry Press, Beijing.

Li, H., & Zhang, W. (1999). Ideal models of overprint of primary halo for large, mega-size blind Au ore deposits. Contributions to Geology and Mineral Resources Research 14 (3), 25-33. https://doi.org/10.6053/j.issn.1001-1412.1999.3.

Li, H., Wang, Z., & Li, F. (1995). Ideal models of superimposed primary halos in hydrothermal gold deposits. Journal of Geochemical Exploration, 55(1-3), 329-336. https://doi.org/10.1016/0375-6742(94)00063-8.

Li, H., Yu, B., Zhang, G., Ma, J., & Li, D. (2016). The development and innovation of blind ore geochemical prospecting method from primary halo and superimposed primary halo to structural superimposed halo. Contributions to Geology and Mineral Resources Research, 31 (1), 92-98.

Li, X. F., Chen, Z. B., & Chen, Z. Q. (2014). Distribution and fractionation of rare earth elements in soil–water system and human blood and hair from a mining area in southwest Fujian Province, China. Environmental Earth Sciences, 72(9), 3599-3608. https://doi.org/10.1007/s12665-014-3271-0.

Li, Y., Hernandez, J. H., Aviles, M., Knappett, P. S., Giardino, J. R., Miranda, R., ... & Morales, J. (2020). Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato, México. Journal of Hydrology, 582, 124517.

Lima, A., De Vivo, B., Cicchella, D., Cortini, M., & Albanese, S. (2003). Multifractal IDW interpolation and fractal filtering method in environmental studies: an application on regional stream sediments of (Italy), Campania region. Applied Geochemistry, 18(12), 1853-1865. https://doi.org/10.1016/S0883-2927(03)00083-0.

Lima, A., Plant, J. A., De Vivo, B., Tarvainen, T., Albanese, S., & Cicchella, D. (2008). Interpolation methods for geochemical maps: a comparative study using arsenic data from European stream waters. Geochemistry: Exploration, Environment, Analysis, 8(1), 41-48. https://doi.org/10.1144/1467-7873/07-146.

Lindner, U., Lingott, J., Richter, S., Jakubowski, N., & Panne, U. (2013). Speciation of gadolinium in surface water samples and plants by hydrophilic interaction chromatography hyphenated with inductively coupled plasma mass spectrometry. Analytical and Bioanalytical Chemistry, 405(6), 1865-1873. https://doi.org/10.1007/s00216-012-6643-x.

Lingott, J., Lindner, U., Telgmann, L., Esteban-Fernández, D., Jakubowski, N., & Panne, U. (2016). Gadolinium-uptake by aquatic and terrestrial organisms-distribution determined by laser ablation inductively coupled plasma mass spectrometry. Environmental Science: Processes & Impacts, 18(2), 200-207. https://doi.org/10.1039/C5EM00533G.

Liu, C., Ma, S., & Hu, Q., (2010). A study of primary halo exploration indexes for metallic ore deposit. Geophysical and Geochemical Exploration, 34 (6), 765-771.

Liu, Y., Cheng, Q., Zhou, K., Xia, Q., & Wang, X. (2016). Multivariate analysis for geochemical process identification using stream sediment geochemical data: A perspective from compositional data. Geochemical Journal, 50(4), 293-314. https://doi.org/10.2343/geochemj.2.0415.

Liu, Y., Zhou, K., & Carranza, E. J. M. (2018). Compositional balance analysis for geochemical pattern recognition and anomaly mapping in the western Junggar region, China. Geochemistry: Exploration, Environment, Analysis, 18(3), 263-276. https://doi.org/10.1144/geochem2017-050.

Lloyd, C. D., Pawlowsky-Glahn, V., & Egozcue, J. J. (2012). Compositional data analysis in population studies. Annals of the Association of American Geographers, 102(6), 1251-1266. https://doi.org/10.1080/00045608.2011.652855.

Lohrke, J., Frisk, A. L., Frenzel, T., Schöckel, L., Rosenbruch, M., Jost, G., ... & Pietsch, H. (2017). Histology and gadolinium distribution in the rodent brain after the administration of cumulative high doses of linear and macrocyclic gadolinium-based contrast agents. Investigative Radiology, 52(6), 324-333. https://dx.doi.org/10.1097/RLI.0000000000000344.

Longo, M. L. (2019). How memory can reduce the vulnerability to disasters: the bradyseism of Pozzuoli in southern Italy. AIMS Geosciences, 5(3), 631-644. https://doi.org/10.3934/geosci.2019.3.631.

Luo, X., Wen, M., & Ou, Y. (2007). Exploration Geochemistry. Metallurgical Industry Press, Beijing.

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., SalibianBarrera, M., Verbeke, T., Koller, M., Conceicao, E.L., & Anna di Palma, M. (2020). Robustbase: Basic Robust Statistics. R Package Version 0, pp. 93-96. http://robustbase.r-forge.r-project.org/.

Marckmann, P., Skov, L., Rossen, K., Heaf, J. G., & Thomsen, H. S. (2007). Case-control study of gadodiamide-related nephrogenic systemic fibrosis. Nephrology Dialysis Transplantation, 22(11), 3174-3178. https://doi.org/10.1093/ndt/gfm261.

Maronna, R., Martin, R.D., & Yohai, V. J. (2006). Robust Statistics: Theory and Methods. John Wiley, New York.

Martín-Fernández, J. A., & Thió-Henestrosa, S., (2016). Compositional Data Analysis. Cham: Springer International Publishing, p. 209. https://doi.org/10.1007/978-3-319-44811-4.

Marty, B., & Tolstikhin, I. N. (1998). CO_2 fluxes from mid-ocean ridges, arcs and plumes. Chemical Geology, 145(3-4), 233-248. https://doi.org/10.1016/S0009-2541(97)00145-9.

Masri, S., LeBrón, A. M., Logue, M. D., Valencia, E., Ruiz, A., Reyes, A., & Wu, J. (2021). Risk assessment of soil heavy metal contamination at the census tract level in the city of Santa Ana, CA: implications for health and environmental justice. Environmental Science: Processes & Impacts, 23, 812-830. https://doi.org/10.1039/D1EM00007A.

McDonald, R. J., McDonald, J. S., Dai, D., Schroeder, D., Jentoft, M. E., Murray, D. L., ... & Kallmes, D. F. (2017). Comparison of gadolinium concentrations within multiple rat organs after intravenous administration of linear versus macrocyclic gadolinium chelates. Radiology, 285(2), 536-545. https://doi.org/10.1148/radiol.2017161594.

McKinley, J. M., Hron, K., Grunsky, E. C., Reimann, C., de Caritat, P., Filzmoser, P., ... & Tolosana-Delgado, R. (2016). The single component geochemical map: Fact or fiction?. Journal of Geochemical Exploration, 162, 16-28. https://doi.org/10.1016/j.gexplo.2015.12.005.

McNutt, S. R., Roman, D. C. (2015). Volcanic seismicity. In: The Encyclopedia of Volcanoes. Academic Press, pp. 1011-1034. https://doi.org/10.1016/B978-0-12-385938-9.00059-6.

McQuitty, L. L. (1966). Similarity analysis by reciprocal pairs for discrete and continuous data. Educational and Psychological Measurement, 26(4), 825-831. https://doi.org/10.1177%2F001316446602600402.

Mebius, L. (1960). A rapid method for the determination of organic carbon in soil. Analytica Chimica Acta, 22, 120-124. https://doi.org/10.1016/S0003-2670(00)88254-9.

Migaszewski, Z. M., & Gałuszka, A. (2021). Abundance and fate of thallium and its stable isotopes in the environment. Reviews in Environmental Science and Bio/Technology, 20, 5-30. https://doi.org/10.1007/s11157-020-09564-8.

Moeck, I. S. (2014). Catalog of geothermal play types based on geologic controls. Renewable and Sustainable Energy Reviews, 37, 867-882. https://doi.org/10.1016/j.rser.2014.05.032.

Molisso, F., Caccavale, M., Capodanno, M., Di Gregorio, C., Gilardi, M., Guarino, A., ... & Sacchi, M. (2020). Sedimentological analysis of marine deposits off the Bagnoli-Coroglio Site of National Interest (SNI), Pozzuoli (Napoli) Bay. Chemistry and Ecology, 36(6), 565-578. https://doi.org/10.1080/02757540.2020.1747447.

Möller, P., Dulski, P., Bau, M., Knappe, A., Pekdeger, A., & Sommer-von Jarmersted, C. (2000). Anthropogenic gadolinium as a conservative tracer in hydrology. Journal of Geochemical Exploration, 69, 409-414. https://doi.org/10.1016/S0375-6742(00)00083-2.

Möller, P., Paces, T., Dulski, P., & Morteani, G. (2002). Anthropogenic Gd in surface water, drainage system, and the water supply of the city of Prague, Czech Republic. Environmental Science & Technology, 36(11), 2387-2394. https://doi.org/10.1021/es010235q.

Moretti, R., De Natale, G., & Troise, C. (2017). A geochemical and geophysical reappraisal to the significance of the recent unrest at Campi Flegrei caldera (Southern Italy). Geochemistry, Geophysics, Geosystems, 18(3), 1244-1269. https://doi.org/10.1002/2016GC006569.

Morroni, L., d'Errico, G., Sacchi, M., Molisso, F., Armiento, G., Chiavarini, S., ... & Regoli, F. (2020). Integrated characterization and risk management of marine sediments: The case study of the industrialized Bagnoli area (Naples, Italy). Marine Environmental Research, 160, 104984. https://doi.org/10.1016/j.marenvres.2020.104984.

Murata, N., Gonzalez-Cuyar, L. F., Murata, K., Fligner, C., Dills, R., Hippe, D., & Maravilla, K. R. (2016). Macrocyclic and other non–group 1 gadolinium contrast agents deposit low levels of gadolinium in brain and bone tissue: preliminary results from 9 patients with normal renal function. Investigative Radiology, 51(7), 447-453. https://doi.org/10.1097/RLI.0000000000000252.

Nasta, P., Palladino, M., Sica, B., Pizzolante, A., Trifuoggi, M., Toscanesi, M., ... & Romano, N. (2020). Evaluating pedotransfer functions for predicting soil bulk density using hierarchical mapping information in Campania, Italy. Geoderma Regional, 21, e00267. https://doi.org/10.1016/j.geodrs.2020.e00267.

Natale, J., Ferranti, L., Isaia, R., Marino, C., Sacchi, M., Spiess, V., ... & Vitale, S. (2021). Integrated on-land-offshore stratigraphy of the Campi Flegrei caldera: new insights into the volcano-tectonic evolution in the last 15 kyr. Basin Research. 1-28. https://doi.org/10.1111/bre.12643.

National Research Council. (2002). Bioavailability of contaminants in soils and sediments: processes, tools and applications. National Research Council (NRC). National Academies: Washington, DC.

Neri, M., Ferrera, E., Giammanco, S., Currenti, G., Cirrincione, R., Patanè, G., & Zanon, V. (2016). Soil radon measurements as a potential tracer of tectonic and volcanic activity. Scientific Reports, 6(1), 1-12. https://doi.org/10.1038/srep24581.

Nozaki, Y., Lerche, D., Alibo, D. S., & Tsutsumi, M. (2000). Dissolved indium and rare earth elements in three Japanese rivers and Tokyo Bay: evidence for anthropogenic Gd and In. Geochimica et Cosmochimica Acta, 64(23), 3975-3982. https://doi.org/10.1016/S0016-7037(00)00472-5.

OECD (2018). Magnetic Resonance Imaging (MRI) Exams. Available online: https://doi.org/10.1787/1d89353f-en (Accessed on 7 December 2018).

Oh, K. Y., Roberts, V. H., Schabel, M. C., Grove, K. L., Woods, M., & Frias, A. E. (2015). Gadolinium chelate contrast material in pregnancy: fetal biodistribution in the nonhuman primate. Radiology, 276(1), 110-118. https://doi.org/10.1148/radiol.15141488.

Orsi, G., De Vita, S., & Di Vito, M. (1996). The restless, resurgent Campi Flegrei nested caldera (Italy): constraints on its evolution and configuration. Journal of Volcanology and Geothermal Research, 74(3-4), 179-214. https://doi.org/10.1016/S0377-0273(96)00063-7.

Padulano, R., Giudice, G. D., Giugni, M., Fontana, N., & Uberti, G. S. D. (2018). Identification of annual water demand patterns in the City of Naples. Proceedings, 2(11). https://doi.org/10.3390/proceedings2110587.

Palarea-Albaladejo, J., & Martín-Fernández, J. A. (2008). A modified EM alr-algorithm for replacing rounded zeros in compositional data sets. Computers & Geosciences, 34(8), 902-917. https://doi.org/10.1016/j.cageo.2007.09.015.

Panthi, S. R., & Wareham, D. G. (2011). Removal of arsenic from water using the adsorbent: New Zealand iron-sand. Journal of Environmental Science and Health, Part A, 46(13), 1533-1538. https://doi.org/10.1080/10934529.2011.609376.

Parsapoor, A., Khalili, M., & Maghami, M. (2017). Discrimination between mineralized and unmineralized alteration zones using primary geochemical haloes in the Darreh-Zar porphyry copper deposit in Kerman, southeastern Iran. Journal of African Earth Sciences, 132, 109-126. https://doi.org/10.1016/j.jafrearsci.2017.04.027.

Patanè, D., De Gori, P., Chiarabba, C., & Bonaccorso, A. (2003). Magma ascent and the pressurization of Mount Etna's volcanic system. Science, 299(5615), 2061-2063. https://doi.org/10.1126/science.1080653.

Pawlowsky-Glahn, V., & Buccianti, A. (2011). Compositional Data Analysis: Theory and Applications. John Wiley & Sons, Chichester.

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. https://doi.org/10.1098/rspl.1896.0076.

Peccerillo, A. (2020). Campania volcanoes: petrology, geochemistry, and geodynamic significance. In: De Vivo, B., Belkin, H.E., Rolandi, G., (Eds), Vesuvius, Campi Flegrei, and Campanian volcanism, Elsevier Amsterdam. pp. 79-120. https://doi.org/10.1016/B978-0-12-816454-9.00005-5.

Peccerillo, A. (1985). Roman comagmatic province (central Italy): evidence for subduction-related magma genesis. Geology, 13(2), 103-106. https://doi.org/10.1130/0091-7613(1985)13%3C103:RCPCIE%3E2.0.CO;2.

Pedreira, R. M., Pahnke, K., Böning, P., & Hatje, V. (2018). Tracking hospital effluent-derived gadolinium in Atlantic coastal waters off Brazil. Water Research, 145, 62-72. https://doi.org/10.1016/j.watres.2018.08.005.

Pedone, M., Aiuppa, A., Giudice, G., Grassa, F., Cardellini, C., Chiodini, G., & Valenza, M. (2014). Volcanic CO_2 flux measurement at Campi Flegrei by tunable diode laser absorption spectroscopy. Bulletin of Volcanology, 76(4), 1-13. https://doi.org/10.1007/s00445-014-0812-z.

Peng, R. D. (2019). simpleboot: Simple Bootstrap Routines. R package version 1.1-7. https://CRAN.R-project.org/package=simpleboot.

Perelman, L., & Ostfeld, A. (2011). Topological clustering for water distribution systems analysis. Environmental Modelling & Software, 26(7), 969-972. https://doi.org/10.1016/j.envsoft.2011.01.006.

Perrotta, A., Scarpati, C., Luongo, G., & Morra, V. (2006). The Campi Flegrei caldera boundary in the city of Naples. In De Vivo, B. (Eds.). Volcanism in the Campania Plain: Vesuvius, Campi Flegrei and Ignimbrites. Elsevier, Amsterdam, pp. 85-96.

Petrik, A., Albanese, S., Lima, A., & De Vivo, B. (2018a). The spatial pattern of beryllium and its possible origin using compositional data analysis on a high-density topsoil data set from the Campania Region (Italy). Applied Geochemistry, 91, 162-173. https://doi.org/10.1016/j.apgeochem.2018.02.008.

Petrik, A., Thiombane, M., Albanese, S., Lima, A., & De Vivo, B. (2018b). Source patterns of Zn, Pb, Cr and Ni potentially toxic elements (PTEs) through a compositional discrimination analysis: A case study on the Campanian topsoil data. Geoderma, 331, 87-99. https://doi.org/10.1016/j.geoderma.2018.06.019.

Petrosino, P., Sadeghi, M., Albanese, S., Andersson, M., Lima, A., & De Vivo, B. (2013). REE contents in solid sample media and stream water from different geological contexts: Comparison between Italy and Sweden. Journal of Geochemical Exploration, 133, 176-201. https://doi.org/10.1016/j.gexplo.2012.12.008.

Petruzzelli, G., Pedron, F., & Rosellini, I. (2020). Bioavailability and bioaccessibility in soil: a short review and a case study. AIMS Environmental Science, 7(2), 208-225. https://doi.org/10.3934/environsci.2020013.

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. https://doi.org/10.1007/s00531-013-0972-7.

Piper, A. M. (1944). A graphic procedure in the geochemical interpretation of water-analyses. Eos, Transactions American Geophysical Union, 25(6), 914-928. https://doi.org/10.1029/TR025i006p00914.

Pizzolante, A., Nicodemo, F., Pierri, A., Ferro, A., Pierri, B., Buonerba, C., ... & Cerino, P. (2021). Development of a municipality index of environmental pressure in Campania, Italy. Future Science OA, 7, FSO720. https://doi.org/10.2144/fsoa-2021-0055.

Prajapati, K., & Modi, H. A. (2012). The importance of potassium in plant growth–a review. Indian Journal of Plant Sciences, 1(02-03), 177-186.

R Core Team (2021). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rabiet, M., Brissaud, F., Seidel, J. L., Pistre, S., & Elbaz-Poulichet, F. (2009). Positive gadolinium anomalies in wastewater treatment plant effluents and aquatic environment in the Hérault watershed (South France). Chemosphere, 75(8), 1057-1064. https://doi.org/10.1016/j.chemosphere.2009.01.036.

Rader, S. T., Maier, R. M., Barton, M. D., & Mazdab, F. K. (2019). Uptake and fractionation of Thallium by Brassica juncea in a geogenic Thallium-amended substrate. Environmental Science & Technology, 53(5), 2441-2449. https://doi.org/10.1021/acs.est.8b06222.

Razum, I., Miko, S., Ilijanić, N., Hasan, O., Miko, M. Š., Brunović, D., & Pawlowsky-Glahn, V. (2020). A compositional approach to the reconstruction of geochemical processes involved in the evolution of Holocene marine flooded coastal karst basins (Mljet Island, Croatia). Applied Geochemistry, 116, 104574. https://doi.org/10.1016/j.apgeochem.2020.104574.

Reimann, C., & De Caritat, P. (1998). Chemical Elements in the Environment: Factsheets for the Geochemist and Environmental Scientist; Springer-Verlag: Berlin/Heidelberg, Germany.

Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A., ... & Team, T. G. P. (2012). The concept of compositional data analysis in practice—total major element concentrations in agricultural and grazing land soils of Europe. Science of the Total Environment, 426, 196-210. https://doi.org/10.1016/j.scitotenv.2012.02.032.

Reimann, C., Filzmoser, P., Garrett, R. & Dutter, R. (2008). Statistical Data Analysis Explained: Applied Environmental Statistics with R. John Wiley & Sons. West Sussex, England. https://doi.org/10.1002/9780470987605.

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. https://doi.org/10.1016/j.scitotenv.2017.06.063.

Rieuwerts, J. S., Thornton, I., Farago, M. E., & Ashmore, M. R. (1998). Factors influencing metal bioavailability in soils: preliminary investigations for the development of a critical loads approach for metals. Chemical Speciation & Bioavailability, 10(2), 61-75. https://doi.org/10.3184/095422998782775835.

Roberts, D. R., Lindhorst, S. M., Welsh, C. T., Maravilla, K. R., Herring, M. N., Braun, K. A., ... & Davis, W. C. (2016). High levels of gadolinium deposition in the skin of a patient with normal renal function. Investigative Radiology, 51(5), 280-289. https://doi.org/10.1097/RLI.0000000000000266.

Rogosnitzky, M., & Branch, S. (2016). Gadolinium-based contrast agent toxicity: a review of known and proposed mechanisms. Biometals, 29(3), 365-376. https://doi.org/10.1007/s10534-016-9931-7.

Rolandi, G., Bellucci, F., Heizler, M. T., Belkin, H. E., & De Vivo, B. (2003). Tectonic controls on the genesis of ignimbrites from the Campanian Volcanic Zone, southern Italy. Mineralogy and Petrology, 79(1), 3-31. https://doi.org/10.1007/s00710-003-0014-4.

Rolandi, G., De Natale, G., Kilburn, C.R., Troise, C., Somma, R., Di Lascio, M., Fedele, A. & Rolandi, R., (2020). The 39 ka Campanian Ignimbrite eruption: new data on source area in the Campanian Plain. Vesuvius, Campi Flegrei, and Campanian Volcanism. Elsevier, p. 175-205. https://doi.org/10.1016/b978-0-12-816454-9.00008-0.

Romano, E., Ausili, A., Zharova, N., Magno, M. C., Pavoni, B., & Gabellini, M. (2004). Marine sediment contamination of an industrial site at Port of Bagnoli, Gulf of Naples, Southern Italy. Marine Pollution Bulletin, 49(5-6), 487-495. https://doi.org/10.1016/j.marpolbul.2004.03.014.

Romano, E., Bergamin, L., Ausili, A., Pierfranceschi, G., Maggi, C., Sesta, G., & Gabellini, M. (2009). The impact of the Bagnoli industrial site (Naples, Italy) on sea-bottom environment. Chemical and textural features of sediments and the related response of benthic foraminifera. Marine Pollution Bulletin, 59(8-12), 245-256. https://doi.org/10.1016/j.marpolbul.2009.09.017.

Romano, E., Bergamin, L., Magno, M. C., Pierfranceschi, G., & Ausili, A. (2018). Temporal changes of metal and trace element contamination in marine sediments due to a steel plant: the case study of Bagnoli (Naples, Italy). Applied Geochemistry, 88, 85-94. https://doi.org/10.1016/j.apgeochem.2017.05.012.

Romano, E., & Gabellini, M. (2001). Analisi delle caratteristiche granulometriche – sedimenti, scheda 3. In: Cicero, A.M., Di Girolamo, I. (Eds.), Metodologie Analitiche di Riferimento del Programma di Monitoraggio per il controllo dell’ambiente marino costiero (triennio 2001-2003). Ministero dell’Ambiente e della Tutela del Territorio, ICRAM, Roma, p. 122.

Romano, P., Santo, A., & Voltaggio, M. (1994). L’evoluzione geo-morfologica della Pianura del Fiume Volturno (Campania) durante il tardo Quaternario (Pleistocene medio superiore-Olocene). Il Quaternario, 7(1), 41-56.

Rosi, M., & Sbrana, A. (1987). Phlegrean Fields. Quaderni De La Ricerca Scientifica, 9(114), 175 pp.

Rossman, L. A. (2000). EPANET 2 User Manual, U.S. Environmental Protection Agency (USEPA), Washington, D.C., EPA/600/R-00/057.

Rout, G., Samantaray, S., & Das, P. (2001). Aluminium toxicity in plants: a review. Agronomie, 21(1), 3-21. https://dx.doi.org/10.1051/agro:2001105.

Ryu, H., Chung, J. S., Nam, T., Moon, H. S., & Nam, K. (2010). Incorporation of heavy metals bioavailability into risk characterization. CLEAN–Soil, Air, Water, 38(9), 812-815. https://doi.org/10.1002/clen.201000070.

Sabbarese, C., Ambrosino, F., Chiodini, G., Giudicepietro, F., Macedonio, G., Caliro, S., ... & Roca, V. (2020). Continuous radon monitoring during seven years of volcanic unrest at Campi Flegrei caldera (Italy). Scientific Reports, 10(1), 1-10. https://doi.org/10.1038/s41598-020-66590-w.

Sacchi, M., Pepe, F., Corradino, M., Insinga, D. D., Molisso, F., & Lubritto, C. (2014). The Neapolitan Yellow Tuff caldera offshore the Campi Flegrei: Stratal architecture and kinematic reconstruction during the last 15 ky. Marine Geology, 354, 15-33. https://doi.org/10.1016/j.margeo.2014.04.012.

Sadeghi, B. (2020). Quantification of Uncertainty in Geochemical Anomalies in Mineral Exploration. PhD thesis. University of New South Wales.

Sakoda, A., Ishimori, Y., & Yamaoka, K. (2011). A comprehensive review of radon emanation measurements for mineral, rock, soil, mill tailing and fly ash. Applied Radiation and Isotopes, 69(10), 1422-1435. https://doi.org/10.1016/j.apradiso.2011.06.009.

Sardans, J., & Peñuelas, J. (2021). Potassium control of plant functions: ecological and agricultural implications. Plants, 10(2), 419. https://doi.org/10.3390/plants10020419.

Sawyer, G., McMcarty, D., & Parkin, G. (2003). Chemistry for environmental engineering and science, McGraw Hill Education, New York.

Sbrana, A., Marianelli, P., & Pasquini, G. (2021). The Phlegrean Fields volcanological evolution. Journal of Maps, 17(2), 545-558. https://doi.org/10.1080/17445647.2021.1982033.

Scandone, R., Bellucci, F., Lirer, L., & Rolandi, G. (1991). The structure of the Campanian Plain and the activity of the Neapolitan volcanoes (Italy). Journal of Volcanology and Geothermal Research, 48(1-2), 1-31. https://doi.org/10.1016/0377-0273(91)90030-4.

Scarpati, C., Perrotta, A., Lepore, S., & Calvert, A. (2013). Eruptive history of Neapolitan volcanoes: constraints from ^{40}Ar–^{39}Ar dating. Geological Magazine, 150(3), 412-425. https://doi.org/10.1017/S0016756812000854.

Schlee, J. S. (1973). Atlantic Continental Shelf and Slope of the United States Sediment Texture of the Northeastern Part. U.S. Geological Survey Professional Paper 529-L, p. 64. https://doi.org/10.3133/pp529L.

Schwarzenbach, R. P., Escher, B. I., Fenner, K., Hofstetter, T. B., Johnson, C. A., Von Gunten, U., & Wehrli, B. (2006). The challenge of micropollutants in aquatic systems. Science, 313(5790), 1072-1077. https://doi.org/10.1126/science.1127291.

Sciarrillo, R., Zuzolo, D., Cicchella, D., Iannone, F., Cammino, G., & Guarino, C. (2020). Contamination and ecological risk assessment of the seaport of Naples (Italy): Insights from marine sediments. Journal of Geochemical Exploration, 210, 106449. https://doi.org/10.1016/j.gexplo.2019.106449.

Scibetta, M., Boano, F., Revelli, R., & Ridolfi, L. (2014). Community detection as a tool for district metered areas identification. Procedia Engineering, 70, 1518-1523. https://doi.org/10.1016/j.proeng.2014.02.167.

Scibetta, M., Boano, F., Revelli, R., & Ridolfi, L. (2013). Community detection as a tool for complex pipe network clustering. Europhysics Letters, 103(4), 48001. https://doi.org/10.1209/0295-5075/103/48001.

Segovia, N. (1991). Radon and volcanic activity: recent advances. International Journal of Radiation Applications and Instrumentation. Part D. Nuclear Tracks and Radiation Measurements, 19(1-4), 409-413. https://doi.org/10.1016/1359-0189(91)90227-9.

Segovia, N., & Mena, M. (1999). Soil radon pulses related to the initial phase of volcanic eruptions. Il nuovo cimento C, 22(3-4), 275-279.

Selinus, O. (2013). Essentials of Medical Geology: Revised Edition. Springer Netherlands, p. 805. https://doi.org/10.1007/978-94-007-4375-5.

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. https://doi.org/10.1016/j.gexplo.2019.02.007.

Semelka, R. C., Commander, C. W., Jay, M., Burke, L. M., & Ramalho, M. (2016a). Presumed gadolinium toxicity in subjects with normal renal function: a report of 4 cases. Investigative Radiology, 51(10), 661-665. https://doi.org/10.1097/RLI.0000000000000318.

Semelka, R. C., Ramalho, J., Vakharia, A., AlObaidy, M., Burke, L. M., Jay, M., & Ramalho, M. (2016b). Gadolinium deposition disease: initial description of a disease that has been around for a while. Magnetic Resonance Imaging, 34(10), 1383-1390. https://doi.org/10.1016/j.mri.2016.07.016.

Sempewo, J., Pathirana A., & Vairavamoorthy, K. (2008). Spatial analysis tool for development of leakage control zones from the analogy of distributed computing. Proceedings of the 10th Annual Water Distribution Systems Analysis (WDSA) Conference, Kruger National Park, American Society of Civil Engineers, pp. 676-690. https://doi.org/10.1061/41024(340)57.

Seyoum, A. G., & Tanyimboh, T. T. (2017). Integration of hydraulic and water quality modelling in distribution networks: EPANET-PMX. Water Resources Management, 31(14), 4485-4503. https://doi.org/10.1007/s11269-017-1760-0.

Shao, Y. (1997). Exploration for Hydrothermal Deposit With Primary Halo Prospecting, 1st ed. iv. Geological Publishing House, Beijing, pp. 143.

Sharp, W. E., & Nardi, G. (1987). A study of the heavy-metal pollution in the bottom sediments at Porto Di Bagnoli (Naples), Italy. Journal of Geochemical Exploration, 29(1-3), 31-48. https://doi.org/10.1016/0375-6742(87)90069-0.

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. https://doi.org/10.1016/j.gexplo.2018.03.003.

Sholkovitz, E. R. (1995). The aquatic chemistry of rare earth elements in rivers and estuaries. Aquatic geochemistry, 1(1), 1-34. https://doi.org/10.1007/BF01025229.

Sieber, M. A., Lengsfeld, P., Frenzel, T., Golfier, S., Schmitt-Willich, H., Siegmund, F., ... & Pietsch, H. (2008). Preclinical investigation to compare different gadolinium-based contrast agents regarding their propensity to release gadolinium in vivo and to trigger nephrogenic systemic fibrosis-like lesions. European Radiology, 18(10), 2164-2173. https://doi.org/10.1007/s00330-008-0977-y.

Siegel, F. (2002). Environmental Geochemistry of Potentially Toxic Metals. Springer, Berlin. https://doi.org/10.1007/978-3-662-04739-2.

Sinclair, A. J., & Blackwell, G. H. (2002). Applied mineral inventory estimation. Cambridge University Press. https://doi.org/10.1017/CBO9780511545993.

Smith, V. C., Isaia, R., & Pearce, N. J. G. (2011). Tephrostratigraphy and glass compositions of post-15 kyr Campi Flegrei eruptions: implications for eruption history and chronostratigraphic markers. Quaternary Science Reviews, 30(25-26), 3638-3660.

Smith, C., & Liu, X. M. (2018). Spatial and temporal distribution of rare earth elements in the Neuse River, North Carolina. Chemical Geology, 488, 34-43. https://doi.org/10.1016/j.chemgeo.2018.04.003.

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. https://doi.org/10.1016/j.chemosphere.2021.129955.

Somma, R., Iuliano, S., Matano, F., Molisso, F., Passaro, S., Sacchi, M., ... & De Natale, G. (2016). High-resolution morpho-bathymetry of Pozzuoli Bay, southern Italy. Journal of Maps, 12(2), 222-230. https://doi.org/10.1080/17445647.2014.1001800.

Song, H., Shin, W. J., Ryu, J. S., Shin, H. S., Chung, H., & Lee, K. S. (2017). Anthropogenic rare earth elements and their spatial distributions in the Han River, South Korea. Chemosphere, 172, 155-165. https://doi.org/10.1016/j.chemosphere.2016.12.135.

Sparks, R. S. J. (2003). Forecasting volcanic eruptions. Earth and Planetary Science Letters, 210(1-2), 1-15. https://doi.org/10.1016/S0012-821X(03)00124-9.

Stellato, L., Coda, S., Arienzo, M., De Vita, P., Di Rienzo, B., D'Onofrio, A., ... & 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. https://doi.org/10.3390/w12123463.

Taylor, H. (1997). Cement Chemistry. Thomas Telford Publishing, London, p. 459. https://doi.org/10.1680/cc.25929.

Tedesco, D., Pece, R., & Avino, R. (1996). Radon, pH and temperature monitoring in water wells at Campi Flegrei caldera (southern Italy). Geochemical Journal, 30(2), 131-138.

Templ, M., Hron, K., & Filzmoser, P. (2011). robCompositions: an R-package for robust statistical analysis of compositional data, in Pawlowsky-Glahn, V., Buccianti A. (eds). Compositional Data Analysis: Theory and Applications, Wiley, Chichester, 2011, pp. 341-355.

Ternes, T., & Joss, A. (Eds.). (2007). Human pharmaceuticals, hormones and fragrances. IWA publishing: London.

Terray, L., Gauthier, P. J., Breton, V., Giammanco, S., Sigmarsson, O., Salerno, G., ... & Falvard, A. (2020). Radon activity in volcanic gases of Mt. Etna by passive dosimetry. Journal of Geophysical Research: Solid Earth, 125(9), e2019JB019149. https://doi.org/10.1029/2019JB019149.

Terray, L., Gauthier, P. J., Salerno, G., Caltabiano, T., La Spina, A., Sellitto, P., & Briole, P. (2018). A new degassing model to infer magma dynamics from radioactive disequilibria in volcanic plumes. Geosciences, 8(1), 27. https://doi.org/10.3390/geosciences8010027.

Thakral, C., & Abraham, J. L. (2009). Gadolinium-induced nephrogenic systemic fibrosis is associated with insoluble Gd deposits in tissues: in vivo transmetallation confirmed by microanalysis. Journal of Cutaneous Pathology, 36(12), 1244-1254. https://doi.org/10.1111/j.1600-0560.2009.01283.x.

Thakral, C., Alhariri, J., & Abraham, J. L. (2007). Long-term retention of gadolinium in tissues from nephrogenic systemic fibrosis patient after multiple gadolinium-enhanced MRI scans: case report and implications. Contrast Media & Molecular Imaging, 2(4), 199-205. https://doi.org/10.1002/cmmi.146.

The Risk Assessment Information System (2021). Available online: https://rais.ornl.gov/ (Accessed in June 2021).

Thiombane, M., Di Bonito, M., Albanese, S., Zuzolo, D., Lima, A., & De Vivo, B. (2019). Geogenic versus anthropogenic behaviour and geochemical footprint of Al, Na, K and P in the Campania region (Southern Italy) soils through compositional data analysis and enrichment factor. Geoderma, 335, 12-26. https://doi.org/10.1016/j.geoderma.2018.08.008.

Thiombane, M., Martín-Fernández, J. A., Albanese, S., Lima, A., Doherty, A., & De Vivo, B. (2018a). Exploratory analysis of multi-element geochemical patterns in soil from the Sarno River Basin (Campania region, southern Italy) through compositional data analysis (CODA). Journal of Geochemical Exploration, 195, 110-120. https://doi.org/10.1016/j.gexplo.2018.03.010.

Thiombane, M., Zuzolo, D., Cicchella, D., Albanese, S., Lima, A., Cavaliere, M., & De Vivo, B. (2018b). Soil geochemical follow-up in the Cilento World Heritage Park (Campania, Italy) through exploratory compositional data analysis and CA fractal model. Journal of Geochemical Exploration, 189, 85-99. https://doi.org/10.1016/j.gexplo.2017.06.010.

Thomsen, H. S. (2017). Are the increasing amounts of gadolinium in surface and tap water dangerous? Acta Radiologica, 58, 259-263.

Thor, K. (2019). Calcium—Nutrient and messenger. Frontiers in Plant Science, 10, 440. https://doi.org/10.3389/fpls.2019.00440.

Todd, D. K. (1980). Groundwater hydrology. Wiley, New York.

Tong, R., Yang, X., Su, H., Pan, Y., Zhang, Q., Wang, J., & Long, M. (2018). Levels, sources and probabilistic health risks of polycyclic aromatic hydrocarbons in the agricultural soils from sites neighboring suburban industries in Shanghai. Science of the Total Environment, 616, 1365-1373. https://doi.org/10.1016/j.scitotenv.2017.10.179.

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. https://doi.org/10.1002/2015GL063621.

Trifuoggi, M., Donadio, C., Mangoni, O., Ferrara, L., Bolinesi, F., Nastro, R. A., ... & Arienzo, M. (2017). Distribution and enrichment of trace metals in surface marine sediments in the Gulf of Pozzuoli and off the coast of the brownfield metallurgical site of Ilva of Bagnoli (Campania, Italy). Marine Pollution Bulletin, 124(1), 502-511. https://doi.org/10.1016/j.marpolbul.2017.07.033.

Troise, C., De Natale, G., Schiavone, R., Somma, R., & Moretti, R. (2019). The Campi Flegrei caldera unrest: Discriminating magma intrusions from hydrothermal effects and implications for possible evolution. Earth-Science Reviews, 188, 108-122. https://doi.org/10.1016/j.earscirev.2018.11.007.

Troll, V. R., Hilton, D. R., Jolis, E. M., Chadwick, J. P., Blythe, L. S., Deegan, F. M., ... & Zimmer, M. (2012). Crustal CO_2 liberation during the 2006 eruption and earthquake events at Merapi volcano, Indonesia. Geophysical Research Letters, 39(11), 1-6. https://doi.org/10.1029/2012GL051307.

Turetta, C., Barbaro, E., Capodaglio, G., & Barbante, C. (2017). Dissolved rare earth elements in the central-western sector of the Ross Sea, Southern Ocean: Geochemical tracing of seawater masses. Chemosphere, 183, 444-453. https://doi.org/10.1016/j.chemosphere.2017.05.142.

Tzatchkov, V. G., Alcocer-Yamanaka, V. H., & Bourguett Ortíz, V. (2006). Graph theory based algorithms for water distribution network sectorization projects. The 8th Annual Water Distribution Systems Analysis Symposium, American Society of Civil Engineers, Cincinnati, Ohio, USA. https://doi.org/10.1061/40941(247)172.

US Food and Drug Administration (2015). FDA Drug Safety Communication: FDA Evaluating the Risk of Brain Deposits with Repeated Use of Gadolinium-Based Contrast Agents for Magnetic Resonance Imaging (MRI). US Food and Drug Administration: Silver Spring, MD, USA.

USEPA (2011). Exposure Factors Handbook: 2011 Edition. Office of Research and Development, Washington, DC, EPA/600/R-09/052F.

USEPA (2004). Lead and compounds (inorganic) CASRN 7439-92-1 | DTXSID2024161. Available online: https://iris.epa.gov/static/pdfs/0277_summary.pdf (Accessed on 26.12.2021).

USEPA (2001). Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment. Office of Emergency and Remedial Response. Washington, DC. EPA 540-R-02-002.

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. https://doi.org/10.1016/S0009-2541(02)00196-1.

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. https://doi.org/10.1016/S0377-0273(03)00402-5.

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. https://doi.org/10.1016/S0377-0273(99)00042-6.

Van den Boogaart, K. G., & Tolosana-Delgado, R. (2013). Analyzing Compositional Data with R, Springer, Berlin, Germany. https://doi.org/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. https://doi.org/10.1007/s00531-013-0979-0.

Wang, D. H. (1995). On the metallogenetic mechanism and model of the Ashele volcanogenic massive sulfide copper deposit, Xinjiang Unpubl. Ph.D. thesis, Graduate School, Chinese Academy of Geological Sciences.

Wang, J., Zang, X., Guo, X., & Xie, H., (2007). The improved Gregorian's zoning index calculating method. Journal of Jilin University (Earth Science Edition) (5), 884-888.

Wang, M., Albanese S., Lima A., Cannatelli C., Cosenza A., Lu W., … & Doherty A. (2015). Compositional analysis and pollution impact assessment: A case study in the Gulfs of Naples and Salerno. Estuarine, Coastal and Shelf Science, 160, 22-32. https://doi.org/10.1016/j.ecss.2015.03.031.

Wáng, Y. X. J., Schroeder, J., Siegmund, H., Idée, J. M., Fretellier, N., Jestin-Mayer, G., ... & Morcos, S. K. (2015). Total gadolinium tissue deposition and skin structural findings following the administration of structurally different gadolinium chelates in healthy and ovariectomized female rats. Quantitative Imaging in Medicine and Surgery, 5(4), 534-545. https://dx.doi.org/10.3978%2Fj.issn.2223-4292.2015.05.03.

Wang, W., Zhao, J., & Cheng, Q. (2014). Mapping of Fe mineralization-associated geochemical signatures using logratio transformed stream sediment geochemical data in eastern Tianshan, China. Journal of Geochemical Exploration, 141, 6-14. https://doi.org/10.1016/j.gexplo.2013.11.008.

Washington, H. S. (1906). The Roman Comagmatic Region. Carnegie Institution of Washington. Washington, D.C.

White, G. W., Gibby, W. A., & Tweedle, M. F. (2006). Comparison of Gd (DTPA-BMA)(Omniscan) versus Gd (HP-DO3A)(ProHance) relative to gadolinium retention in human bone tissue by inductively coupled plasma mass spectroscopy. Investigative Radiology, 41(3), 272-278. http://dx.doi.org/10.1097/01.rli.0000186569.32408.95.

Wright, D. B., London, K., & Field, A. P. (2011). Using bootstrap estimation and the plug-in principle for clinical psychology data. Journal of Experimental Psychopathology, 2(2), 252-270. https://doi.org/10.5127%2Fjep.013611.

Wu, Y. F., Yang, F. Q., Liu, F., Zhou, M., & Chen, H. Q. (2015). 40Ar-39Ar dating of sericite from the brittle ductile shear zone in the Ashele Cu-Zn Ore district, Xinjiang. Acta Geoscientica Sinica, (1), 121-126. https://doi.org/10.3975/cagsb.2015.01.15.

Yan, B., & Hou, Y. (2018). Effect of soil magnesium on plants: a review. In IOP Conference Series: Earth and Environmental Science, 170(2), 022168. https://doi.org/10.1088/1755-1315/170/2/022168.

Yang, F., Wu, Y., Yang, J., & Zheng, J., (2016). Metallogenetic model for VMS type polymetallic copper deposits in the Ashele Ore Dense district of Altay, Xinjiang. Geotectonica et Metallogenia, 4, 701-715.

Yang, S., Zhao, J., Chang, S. X., Collins, C., Xu, J., & Liu, X. (2019). Status assessment and probabilistic health risk modeling of metals accumulation in agriculture soils across China: A synthesis. Environment International, 128, 165-174. https://doi.org/10.1016/j.envint.2019.04.044.

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. Chaos: An Interdisciplinary Journal of Nonlinear Science, 21(1), 016111. https://doi.org/10.1063/1.3540339.

Ye, Q., Fu, X., & Wang, B. (1998). Metallogeny of polymetallic metallogenic belts on the southern margin of the Altay Mountains, Xinjiang, China. Acta Geologica Sinica, 72(4), 349-357.

Yruela, I. (2005). Copper in plants. Brazilian Journal of Plant Physiology, 17, 145-156. https://doi.org/10.1590/S1677-04202005000100012.

Zhang, L., McKinley, J., Cooper, M., Han, W., Liu, F., Song, Y., ... & Cheng, H. (2020). Transfer processes of potential toxic elements (PTE) between rock-soil systems and soil risk evaluation in the Baoshan area, Yunnan Province, Southwest China. Applied Geochemistry, 121, 104712. https://doi.org/10.1016/j.apgeochem.2020.104712.

Zhang, Y., Gu, X., Cheng, W., Dong, S., Jing, L., Wang, L., & Zhang, Y. (2010). The geochemical features of primary halo and the evaluation of deep mineralization prospect of Liubagou gold deposit, Inner Mongolia. Earth Science Frontiers, 17(2), 209-221.

Zheng, C., Luo, X., Wen, M., Ebrahimi, P., Liu, P., Liu, G., ... & Albanese, S. (2020). Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China. Journal of Geochemical Exploration, 213, 106509. https://doi.org/10.1016/j.gexplo.2020.106509.

Zhu, Y., Hoshino, M., Yamada, H., Itoh, A., & Haraguchi, H. (2004). Gadolinium anomaly in the distributions of rare earth elements observed for coastal seawater and river waters around Nagoya City. Bulletin of the Chemical Society of Japan, 77(10), 1835-1842. https://doi.org/10.1246/bcsj.77.1835.

Zissimos, A. M., Cohen, D. R., Christoforou, I. C., Sadeghi, B., & Rutherford, N. F. (2021). Controls on soil geochemistry fractal characteristics in Lemesos (Limassol), Cyprus. Journal of Geochemical Exploration, 220, 106682. https://doi.org/10.1016/j.gexplo.2020.106682.

Zuo, R., & Wang, J. (2020). ArcFractal: An ArcGIS add-in for processing geoscience data using fractal/multifractal models. Natural Resources Research, 29(1), 3-12. https://doi.org/10.1007/s11053-019-09513-5.

Zuzolo, D., Cicchella, D., Lima, A., Guagliardi, I., Cerino, P., Pizzolante, A., ... & Albanese, S. (2020). Potentially toxic elements in soils of Campania region (Southern Italy): Combining raw and compositional data. Journal of Geochemical Exploration, 213, 106524. https://doi.org/10.1016/j.gexplo.2020.106524.

APPENDED PUBLICATIONS

Publication No. 1

Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., ... & Albanese, S. (2022). Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy). Journal of Geochemical Exploration, 106922. https://doi.org/10.1016/j.gexplo.2021.106922.

Publication No. 2

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. https://doi.org/10.1039/D0EW01123A.

Publication No. 3

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. https://doi.org/10.1016/j.chemosphere.2021.129955.

Publication No. 4

Zheng, C., Luo, X., Wen, M., Ebrahimi, P., Liu, P., Liu, G., ... & Albanese, S. (2020). Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China. Journal of Geochemical Exploration, 213, 106509. https://doi.org/10.1016/j.gexplo.2020.106509.

Publication No. 5

Ebrahimi, P., & Barbieri, M. (2019). Gadolinium as an emerging microcontaminant in water resources: threats and opportunities. Geosciences, 9(2), 93. https://doi.org/10.3390/geosciences9020093.

Publication No. 1

Ebrahimi, P., Guarino, A., Allocca, V., Caliro, S., Avino, R., Bagnato, E., ... & Albanese, S. (2022). Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy). Journal of Geochemical Exploration, 106922. https://doi.org/10.1016/j.gexplo.2021.106922.

Journal of Geochemical Exploration 233 (2022) 106922

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

journal homepage: www.elsevier.com/locate/gexplo

Hierarchical clustering and compositional data analysis for interpreting groundwater hydrogeochemistry: The application to Campi Flegrei volcanic aquifer (south Italy)

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, Sezione di Napoli, "Osservatorio Vesuviano", Naples 80124, Italy
[c] Department of Earth Sciences, University of Florence, Florence 50121, Italy

ARTICLE INFO

Keywords:
Hierarchical cluster analysis
Groundwater evolution
Hydrothermal system
Precursors of volcanic eruption
Compositional nature of data
Mixing process

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, groundwater samples 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 groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The first group (As = 158.2 ± 169 μg/l, electric conductivity = 1,732.1 ± 1,086 μS/cm and temperature = 25.6 ± 8 °C) is mainly derived from poor arsenic meteoric water, but there is significant thermal/seawater contribution in the second one (As = 1,457.8 ± 2,210 μg/l, electric conductivity = 20,118.3 ± 11,139 μS/cm and temperature = 37.1 ± 20 °C). Interaction of the bicarbonate-rich groundwater and hydrothermal vapors gives rise to the sulfate-rich groundwater (As = 847.2 ± 679 μg/l, electric conductivity = 3,940.0 ± 540 μS/cm and temperature = 82.8 ± 3 °C) around Solfatara volcano. The mixed groundwater (As = 451.4 ± 388 μg/l, electric conductivity = 4,482.9 ± 4,027 μS/cm and temperature = 37.1 ± 16 °C) is observed where the three main groundwater groups undergo a mixing process, depending on the hydrogeology of the volcanic aquifer. Contrary to the bicarbonate- and sulfate-rich groundwater, the chlorine-rich and mixed groundwater generally occurs at low piezometric levels (approximately <1 m above sea level) near the coastline. The hierarchical cluster analysis provides 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 and identifying promising precursors of volcanic eruption.

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

---

[*] Corresponding author.
E-mail addresses: pooria.ebrahimi@unina.it, pooria.ebrahimi@gmail.com (P. Ebrahimi).

https://doi.org/10.1016/j.gexplo.2021.106922
Received 7 June 2021; Received in revised form 7 November 2021; Accepted 26 November 2021

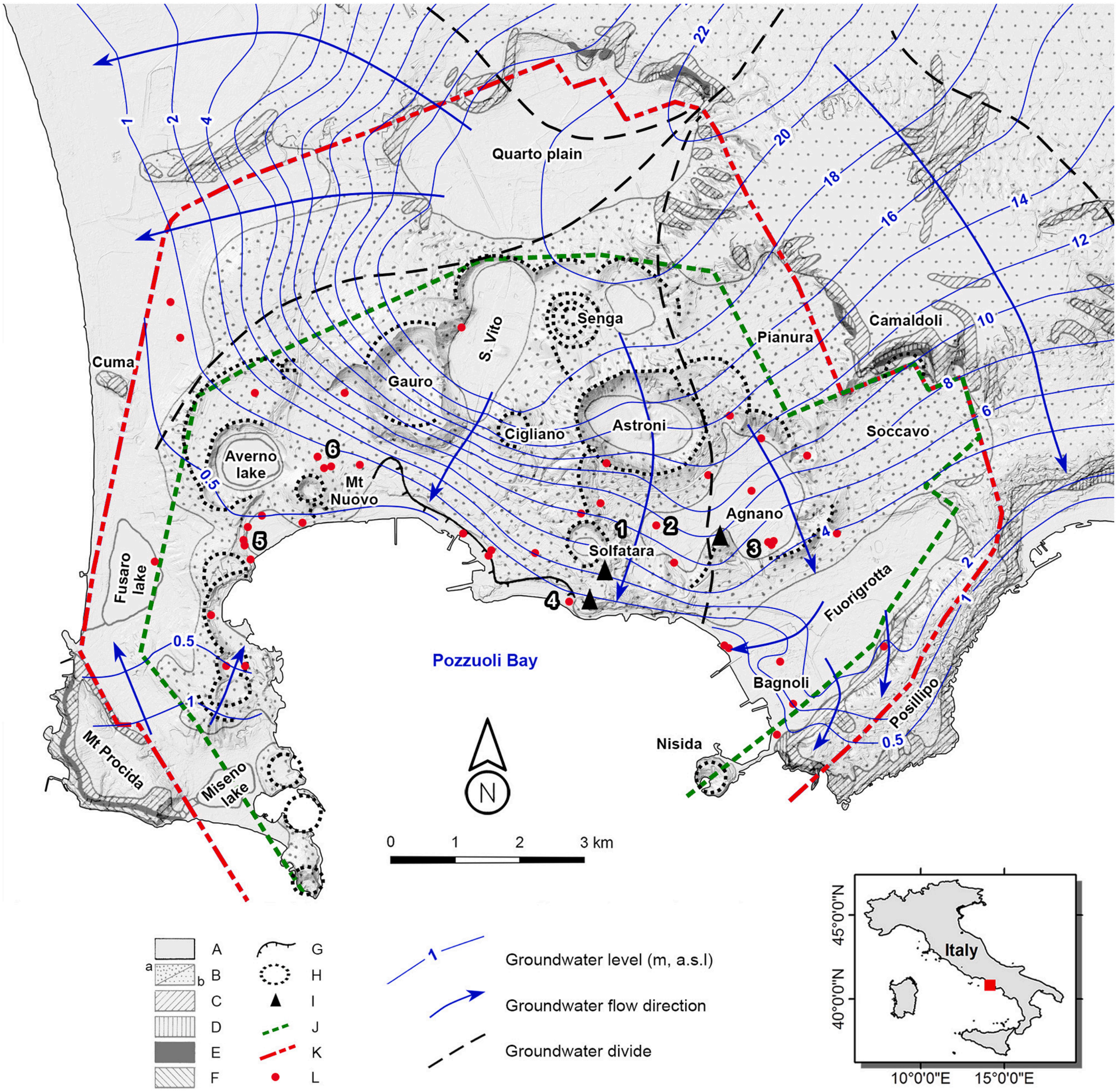
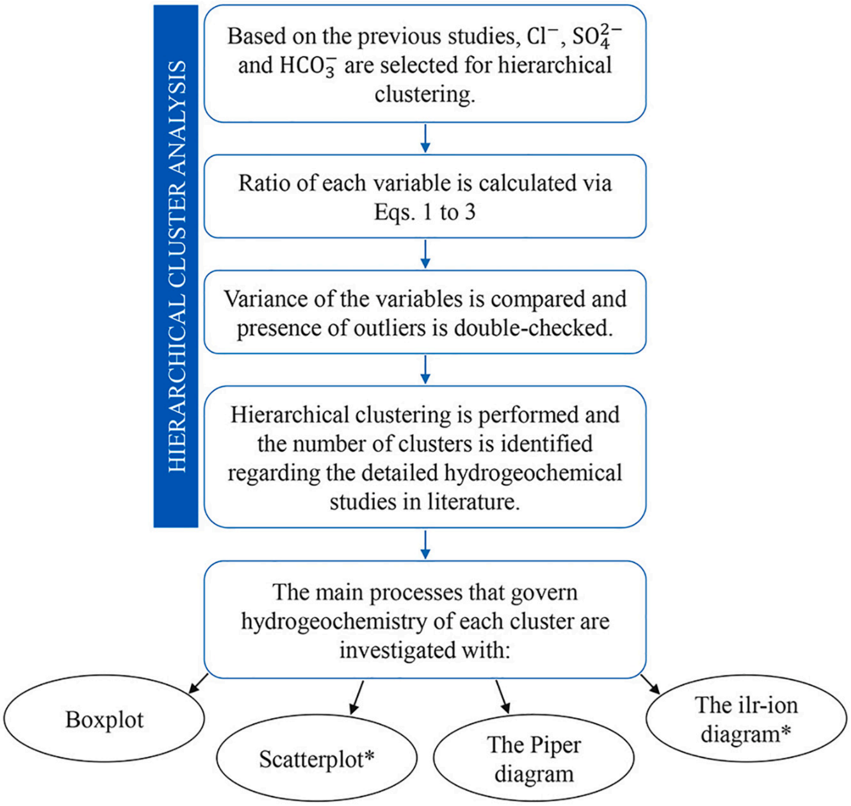


Fig. 1. The simplified geological map of Campi Flegrei (after Valentino et al., 1999). The groundwater level, groundwater flow direction, groundwater divide (after De Vita et al., 2018) together with location of (1) Pisciarelli, (2) Hotel Tennis, (3) Agnano Terme, (4) Terme Puteolane, (5) Stufe di Nerone 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); and (L) sampling points.

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 or bradyseism, 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



* Generated based on the principles of compositional data analysis

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

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 explains only a part of the whole composition. Karl Pearson introduced the problem of spurious correlation in this type of data (Pearson, 1897) 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, 2022; 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 research 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 CE. 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 m 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.,

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

ID	Groundwater 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	1,040.0	817.0	0.7	9.6	423.0	168.1	1,457.5	3,400	86.2	7.0
W2	Na–SO_4	SGW	C4$_{SGW}$	89.2	2.0	1,000.0	288.0	119.0	1,880.0	987.0	<D.L.	3.1	629.0	1,526.3	27,900.0	4,480	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	1,420	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	1,040.0	2,550	36.0	7.7
W5	Na–Cl	ClGW	C3$_{ClGW}$	164.0	11.4	7,740.0	176.0	10,900.0	1,410.0	2,260.0	<D.L.	1.3	11,000.0	3.2	88,812.5	23,000	20.0	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	1,690	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	5,378.8	2,650	44.3	7.3
W8	Na–Cl	ClGW	C2$_{MGW}$	263.0	57.0	2,020.0	327.0	2,880.0	376.0	1,300.0	2.8	2.1	1,230.0	322.5	13,837.5	8,110	56.0	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	1,656.3	2,480	35.9	7.8
W10	Na–Cl	ClGW	C3$_{ClGW}$	564.0	872.0	11,200.0	761.0	20,700.0	2,240.0	807.0	7.9	2.4	4,410.0	3,538.8	23,800.0	37,000	67.0	6.1
W11	Na–Cl	ClGW	C3$_{ClGW}$	382.0	244.0	7,870.0	539.0	14,800.0	317.0	48.8	1.2	6.0	4,210.0	<D.L.	9,927.5	33,100	43.7	6.5
W12	Na–Cl	ClGW	C3$_{ClGW}$	182.0	12.8	3,850.0	250.0	6,330.0	200.0	810.0	65.6	2.7	6,660.0	1,610.0	10,470.0	14,780	29.5	6.5
W13	Na–Cl	ClGW	C2$_{MGW}$	59.6	0.7	1,070.0	378.0	932.0	1,090.0	758.0	0.7	3.8	645.0	968.9	19,825.0	5,160	75.0	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	2,888.8	17,160	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	1,407.5	1,230	–	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	3,170.0	2,310	26.4	5.7
W18	Na–Cl	ClGW	C2$_{MGW}$	424.0	31.4	1,690.0	282.0	2,230.0	368.0	1,190.0	1.7	1.8	329.0	524.6	7,246.3	5,600	21.4	6.4
W19	Na–Cl	ClGW	C3$_{ClGW}$	516.0	526.0	7,110.0	829.0	12,600.0	1,560.0	1,620.0	1.7	3.0	2,610.0	23.2	11,098.8	25,500	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	5,226.3	2,680	23.5	7.6
W21	Na–Cl	ClGW	C2$_{MGW}$	127.0	30.0	843.0	143.0	931.0	134.0	1,150.0	16.0	2.9	264.0	473.8	7,005.0	3,380	36.2	6.9
W22	Na–Cl	ClGW	C3$_{ClGW}$	309.0	60.5	7,370.0	207.0	11,000.0	611.0	371.0	132.0	13.7	11,400.0	6,875.0	31,700.0	26,000	77.0	6.9
W23	Na–Cl	ClGW	C3$_{ClGW}$	168.0	77.9	2,430.0	208.0	3,990.0	405.0	870.0	33.0	6.0	2,380.0	793.6	13,975.0	11,490	32.0	7.0
W24	Na–Cl	ClGW	C3$_{ClGW}$	188.0	29.0	2,070.0	148.0	3,560.0	194.0	336.0	80.9	2.4	2,330.0	262.9	6,111.3	9,400	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	1,021	21.0	7.4
W26	Na–HCO_3	BGW	C1$_{BGW}$	317.0	26.1	498.0	138.0	510.0	613.0	1,000.0	0.4	0.4	363.0	724.4	16,587.5	2,770	40.5	6.2
W27	Na–HCO_3	BGW	C1$_{BGW}$	180.0	68.4	637.0	99.4	551.0	430.0	1,130.0	355.0	3.1	196.0	33.6	1,058.9	3,770	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	1,550	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	1,360	17.4	7.4
W31	Na–HCO_3	BGW	C1$_{BGW}$	162.0	17.4	724.0	134.0	179.0	256.0	1,870.0	0.6	3.1	883.0	154.8	1,776.3	2,490	26.0	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	1,340	22.3	7.6
W33	Na–HCO_3	BGW	C1$_{BGW}$	118.0	29.9	823.0	110.0	412.0	247.0	1,700.0	7.7	2.2	980.0	292.4	4,681.3	3,200	35.7	6.3
W34	Na–HCO_3	BGW	C1$_{BGW}$	149.0	26.7	764.0	136.0	462.0	148.0	1,590.0	0.6	3.2	770.0	166.4	5,713.8	3,060	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	2,432.5	1,218	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	1,666.0	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	1,159	18.8	7.3
W39	Na–HCO_3	BGW	C2$_{MGW}$	39.4	5.1	1,690.0	157.0	1,100.0	530.0	1,940.0	24.1	3.7	928.0	1,590.0	13,437.5	6,340	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	4,825.0	1,878	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	1,510.0	62.6	8.8	386.0	437.0	9,186.3	3,880	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	1,838.8	1,020	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	11,200.0	829.0	20,700.0	2,240.0	2,260.0	355.0	13.7	11,400.0	6,875.0	88,812.5	37,000	86.2	8.6
Average				143.9	51.9	1,567.8	172.8	2,278.7	409.7	757.6	56.3	5.0	2,026.5	561.3	8,260.1	6,469	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	5,090	1,180	10,200	3,160	41	–	–	13,000	11,000	140,000	–	337	6.0
GEW (Mofete 1)[c]				555	–	10,025	1,230	17,710	615	81	–	–	25,000	9,000	125,000	–	250	7.5
GEW (Mofete 1)[c]				1,281	5	12,589	2,342	22,810	670	46	–	–	28,000	11,000	110,000	–	250	6.5
LSW[d]				447	1,397	12,250	446	22,390	2,906	161	–	–	170	2	4,410	–	20	8.1

<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 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-, chlorine- and bicarbonate-rich groundwater, respectively.

[b] The groups are based on hierarchical cluster analysis: C1$_{BGW}$, C2$_{MGW}$, C3$_{ClGW}$ and C4$_{SGW}$ denote the bicarbonate-rich, mixed, chlorine-rich and sulfate-rich groundwater, respectively.

[c] Guglielminetti (1986).

[d] Aiuppa et al. (2006).

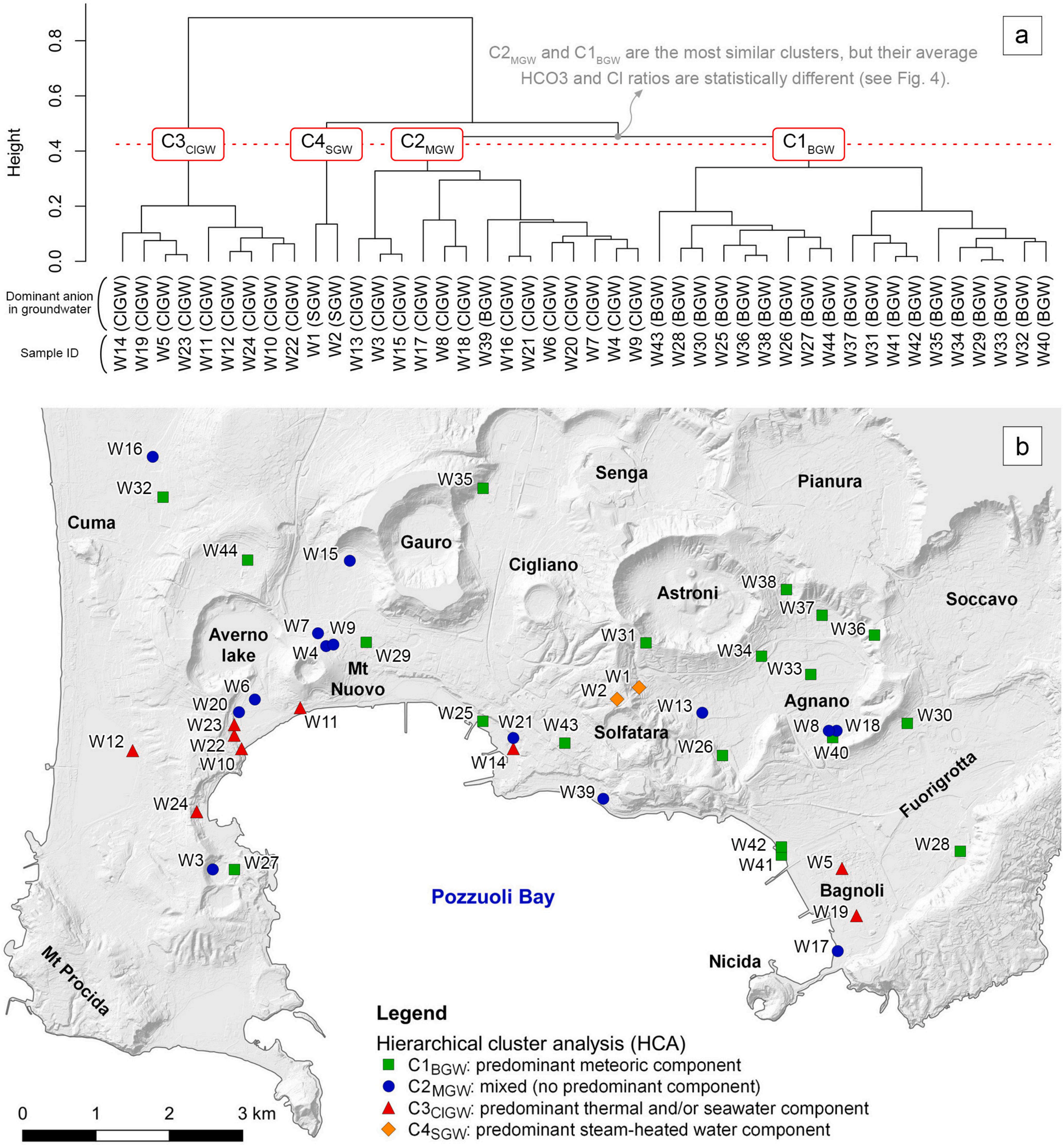


Fig. 3. (a) The dendrogram of hierarchical cluster analysis; and (b) spatial distribution of each group of groundwater samples. The height on the vertical axis of the dendrogram states the distance measure, indicating higher similarity at a lower value. On the horizontal axis of the dendrogram, BGW (bicarbonate-rich groundwater), ClGW (chlorine-rich groundwater) and SGW (sulfate-rich groundwater) refer to the dominant anion group. $C1_{BGW}$: bicarbonate-rich groundwater; $C2_{MGW}$: mixed groundwater; $C3_{ClGW}$: chlorine-rich groundwater; and $C4_{SGW}$: sulfate-rich groundwater.

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).

Little is known about the hydrologic system of Campi Flegrei volcano at the southwestern boundary of Campania Plain. 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

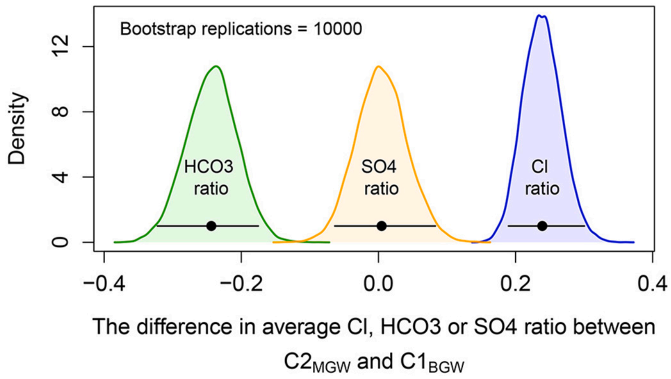


Fig. 4. Density plot of the difference in average HCO_3^-, SO_4^{2-} or Cl^- ratio between the mixed ($C2_{MGW}$) and bicarbonate-rich ($C1_{BGW}$) groundwater. The dots show the values obtained from the dataset in the present article and the horizontal line indicates the uncertainty based on the 95% confidence interval of 10,000 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.

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).

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 <2,000 m) with seawater and local meteoric water origin, and a deep reservoir (depth >2,000 m) containing magmatic-meteoric hypersaline fluids feed the geothermal system of Campi Flegrei (Caprarelli et al., 1997). High As concentration 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 groundwater 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 the 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 water (Stufe di Nerone well and Stufe di Nerone spring); (2) alkali-bicarbonate, less saline water (Agnano spring and the water from Quarto plain); (3) alkali-bicarbonate-sulfate water (a sampling station in Quarto plain); and (4) alkaline earth-sulfate, acidic water (Pisciarelli Grande and Pisciarelli Piccola). They proposed that the groundwater from Pozzanghera Agnano Terme, Hotel Tennis, Terme Puteolane and Damiani is 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 groundwater. The latter classification can be considered as the generalized version of the former, being beneficial when investigating a large number of samples.

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 (EC), pH and temperature are measured with portable instruments in situ. 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

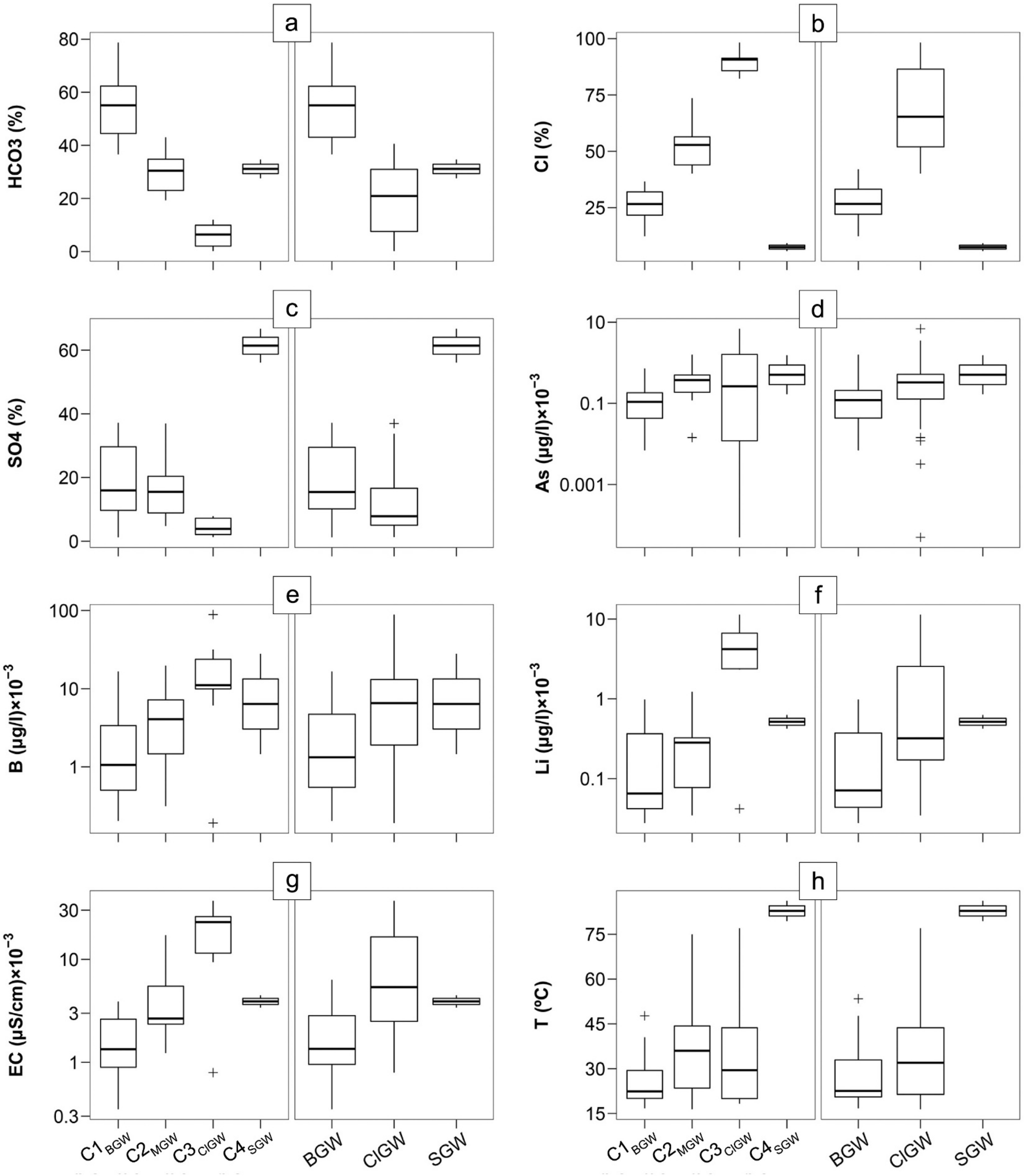


Fig. 5. Variation of selected parameters in the HCA groups and dominant anion groups (the left and right panels in each paired graph, respectively). To generate panels a to c, the following calculations are made using bicarbonate, chloride and sulfate in meq/l: $HCO_3^- = (HCO_3^-/(HCO_3^- + Cl^- + SO_4^{2-})) \times 100$; $Cl^- = (Cl^-/(HCO_3^- + Cl^- + SO_4^{2-})) \times 100$; and $SO_4^{2-} = (SO_4^{2-}/(HCO_3^- + Cl^- + SO_4^{2-})) \times 100$. Scale of y-axis is logarithmic in As, B, Li and EC (d to g, respectively) boxplots. EC stands for electric conductivity and the abbreviations on the horizontal axes are the same with Fig. 3.

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):

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

$$SO_4^{2-} \text{ ratio} = \frac{SO_4^{2-}}{HCO_3^- + SO_4^{2-} + Cl^-} \tag{2}$$

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

where all values are in mg/l. Presence of outliers and variance of the above-mentioned ratios are double-checked prior to hierarchical clus-

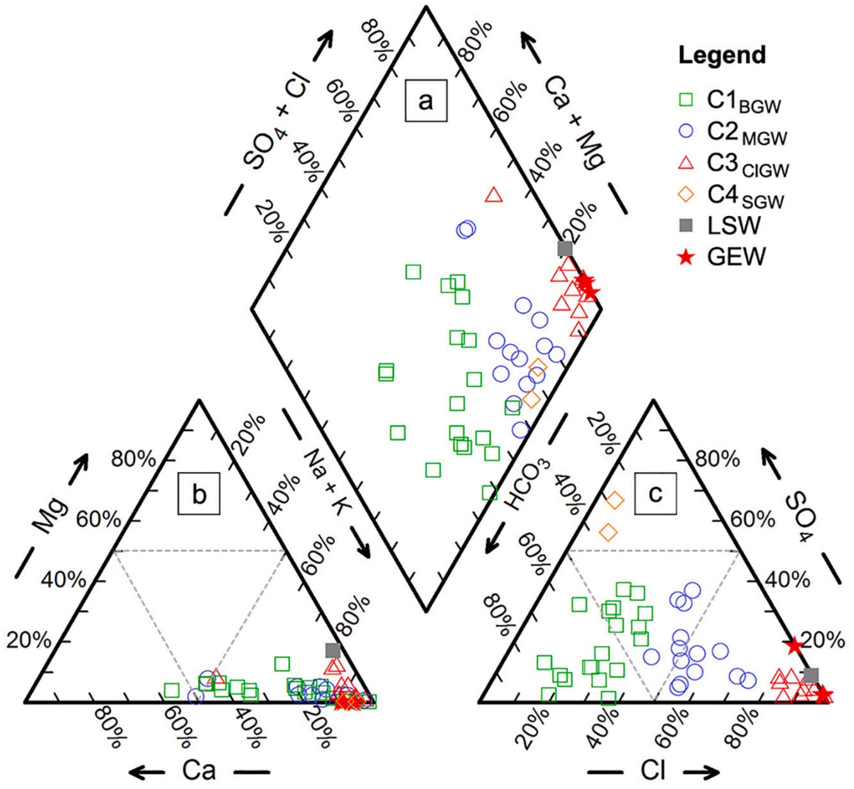


Fig. 6. Piper diagram showing composition of Campi Flegrei groundwater, grouped according to HCA results, together with geothermal water (GEW) and local seawater (LSW): (a) the diamond-shaped field; (b) the cation field; and (c) the anion field. The abbreviations in the legend are the same with Fig. 3.

tering (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, 1897; 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:

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

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

$$z_3 = \sqrt{\frac{2}{3}} ln \frac{\sqrt{[Cl^-][SO_4^{2-}]}}{[HCO_3^- + CO_3^{2-}]} \tag{6}$$

and

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

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

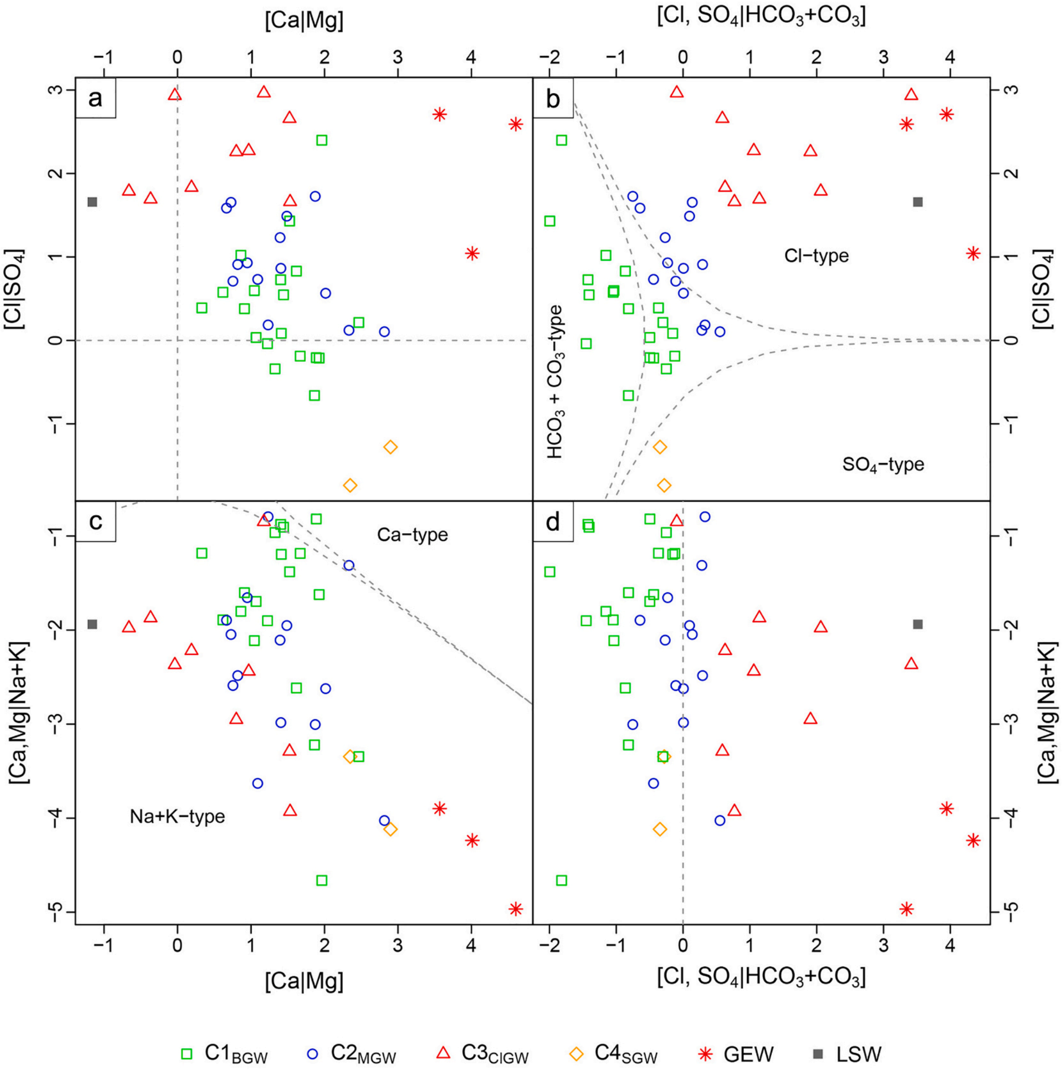


Fig. 7. Chemical composition of the collected groundwater samples on the ilr-ion plot: (a) the information in this panel is not specifically presented in the Piper diagram; (b) corresponds to the anion field of the Piper diagram; (c) corresponds to the cation field of the Piper diagram; and (d) corresponds to the diamond-shaped field of the Piper diagram. The abbreviations are the same with Fig. 3.

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 (e.g., Somma et al., 2021). The R package "robCompositions" (Filzmoser et al., 2018) is used for computing the symmetric coordinates in Figs. 8 and 9. The interested readers are referred to Kynčlová et al. (2017) for the calculation steps. 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^-:

$$\left[\mathrm{Na}^+, \mathrm{Cl}^- | \mathrm{HCO}_3^-\right] = \sqrt{\frac{2}{3}} ln \frac{\sqrt{[\mathrm{Na}^+][\mathrm{Cl}^-]}}{[\mathrm{HCO}_3^-]} \qquad (8)$$

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

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

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 groundwater, respectively) are exclusively used for discussing the dominant anion groups. However, the HCA groups are *mainly* referred to as bicarbonate-rich, mixed, chlorine-rich and sulfate-rich groundwater ($C1_{BGW}$, $C2_{MGW}$, $C3_{ClGW}$ and $C4_{SGW}$, respectively).

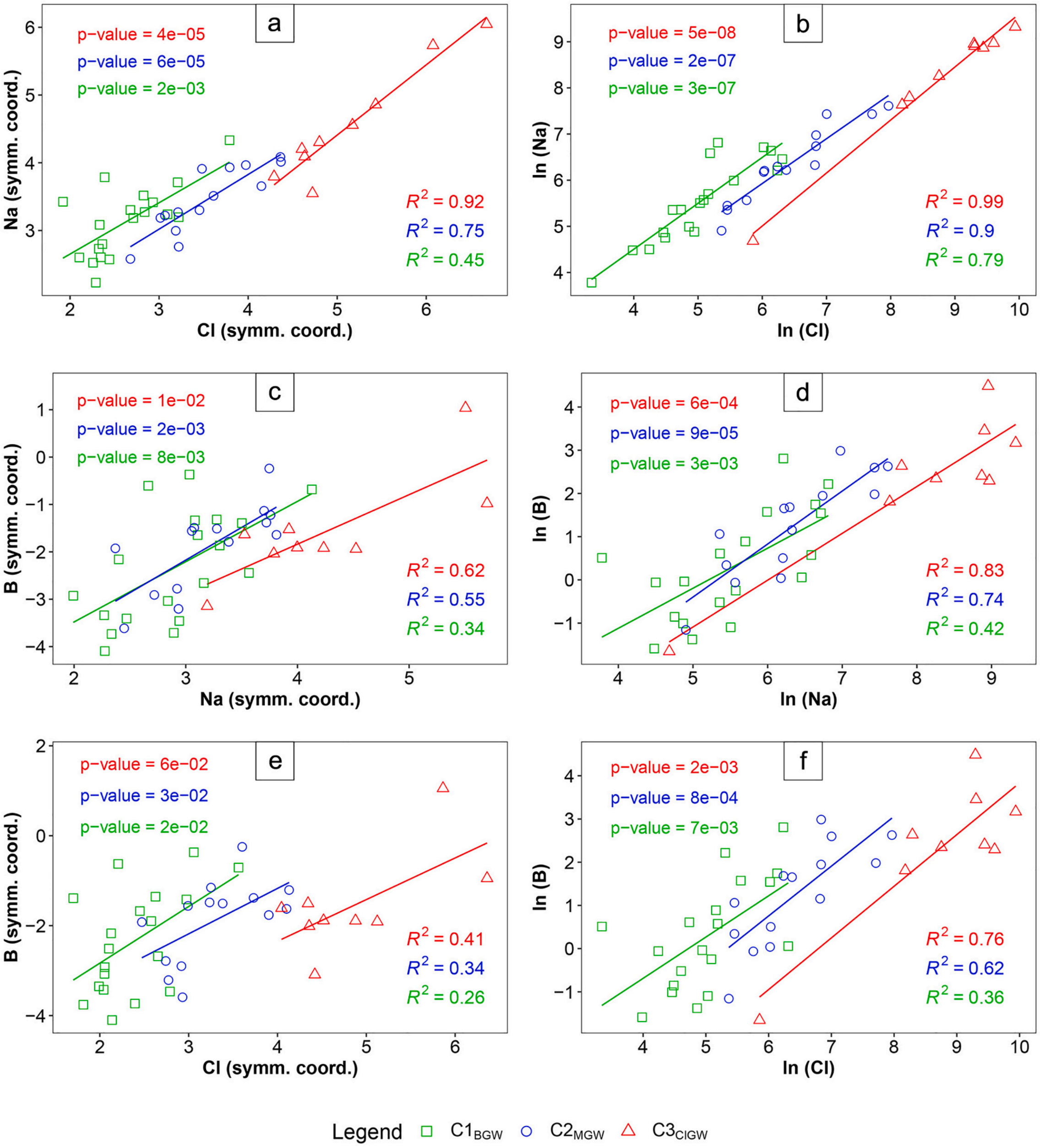


Fig. 8. The relationship between Na^+, Cl^- and B considering the symmetric coordinates (a, c and e) and the classical log-transformed data (b, d and f). The ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-}, Cl^-, NO_3^- and F^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity were used for calculation of the symmetric coordinates. The sulfate-rich groundwater is excluded because of few samples. The abbreviations in the legend are the same with Fig. 3.

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 groundwater around Solfatara crater with the highest temperatures (>79 °C). On average, Na–Cl groundwater is 10 °C warmer than Na–HCO_3 groundwater. The average contents of cations and anions follow the orders of Na^+ (1,567.8 mg/l) > K^+ (172.8 mg/l) > Ca^{2+} (143.9 mg/l) > Mg^{2+} (51.9 mg/l) and Cl^- (2,278.7 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 to 8.6 in Na–HCO_3 groundwater revealing the acidic to alkaline nature of the groundwaters under investigation. Nonetheless, electric conductivity indicates a reverse trend, decreasing from Na–Cl groundwater to Na–HCO_3 groundwater (Table 1).

Regarding the dominant anion groups in Table 1, BGW (T = 16.7 to 53.4 °C and EC = 350 to 6,340 μS/cm) mostly represents the meteoric-derived water which can be supported by the isotopic investigations of Caprarelli et al. (1997). Mixing this groundwater 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 °C and EC = 795 to 37,000 μS/cm). On the other hand, the interaction between BGW and the vapor phase forms SGW with high temperature (>79 °C) and intermediate electric conductivity (3,400 to 4,480 μS/cm). Chiodini et al. (2001, 2003) reported high $H_2S_{(g)}$ in the condensing hydrothermal vapors. Oxidizing H_2S by atmospheric gases results in elevated sulfate concentration in the groundwater around Solfatara crater.

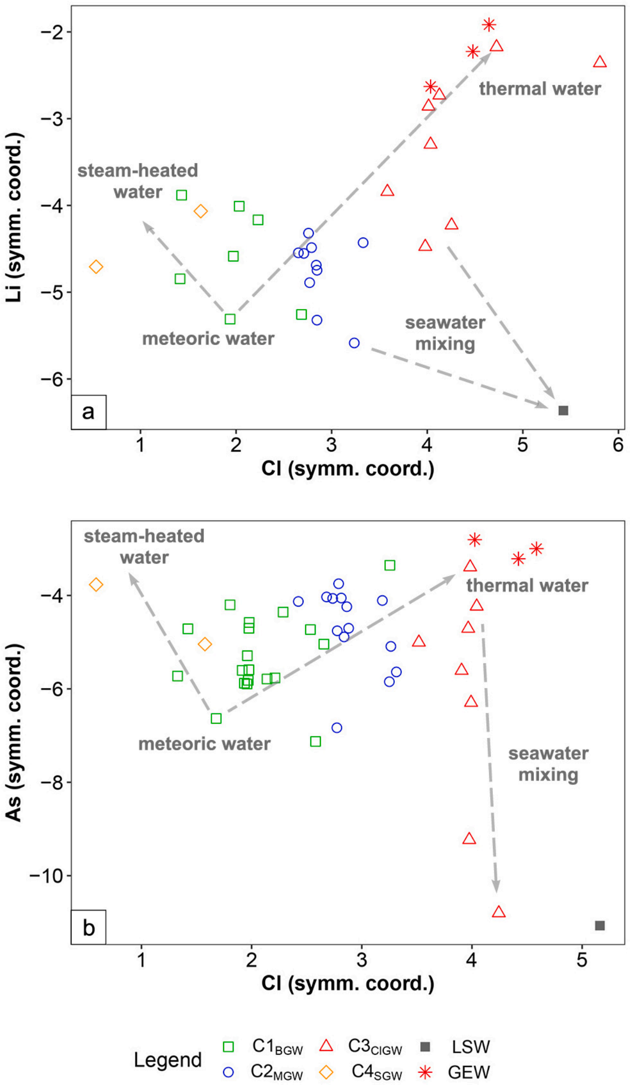


Fig. 9. Li–Cl⁻ and As–Cl⁻ scatterplots (a and b, respectively) generated by the symmetric coordinates. The ions (Ca^{2+}, Mg^{2+}, Na^+, K^+, HCO_3^-, SO_4^{2-} and Cl^-) as the main water components along with As, Li and B as the indicators of hydrothermal activity are used for calculation of the symmetric coordinates. The samples with a lithium content of below detection limit are excluded in the Li–Cl⁻ panel. The abbreviations in the legend are the same with Fig. 3.

4.2. Detection of homogeneous groundwater 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 (Fig. S1a to c). However, computing the anion ratios alleviates these data problems for applying clustering algorithm (Fig. S1d to f). The collected samples can be categorized in the bicarbonate-rich, sulfate-rich, chlorine-rich and mixed groundwater 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, the mixed and chlorine-rich groundwater 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 anion groups, hierarchical cluster analysis divides ClGW into (Fig. 3):

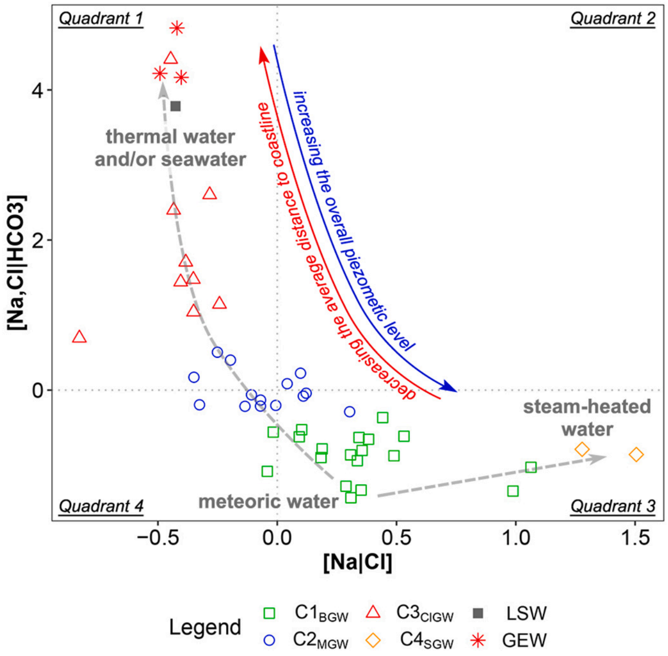


Fig. 10. The relationship between the isometric log-ratios computed using Na^+, Cl^- and HCO_3^- in mg/l. The abbreviations in the legend are the same with Fig. 3.

(1) *mixed groundwater* ($C2_{MGW}$). 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) *chlorine-rich groundwater* ($C3_{ClGW}$). The groundwater around Stufe di Nerone well and Stufe di Nerone spring, and the groundwater 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 the groundwater from Stufe di Nerone well and Stufe di Nerone spring also suggested 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 groundwater 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. 3a, 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 through calculating the distances between observations. Significance of difference in average of the ratios between other combinations of clusters is not assessed due to their few observations and limitation of the bootstrapping approach. The sulfate-rich groundwater is the next group that merges these two at a higher level of the dendrogram (Fig. 3a), probably showing a relatively stronger fingerprint of meteoric water in composition of groundwater in the study area, except for the chlorine-rich groundwater.

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 Fig. 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 Fig. 5b, d, e, f, g and h, respectively) show a reverse trend. Values of the selected parameters in the mixed groundwater range between the corresponding values in bicarbonate- and chlorine-rich groundwater which confirm their hybrid nature. Compared to the bicarbonate-rich groundwater (As = 158.2 ± 169 μg/l, B = 2,875.3 ± 3,973 μg/l and EC = 1,732.1 ± 1,086 μS/cm), the average arsenic, boron and electric conductivity values are almost double in the mixed groundwater (As = 451.4 ± 388 μg/l, B = 5,955.1 ± 5,679 μg/l and EC = 4,482.9 ± 4,027 μS/cm) and one order of magnitude greater in the chlorine-rich groundwater (As = 1,457.8 ± 2,210 μg/l, B = 21,787.4 ± 25,278 μg/l and EC = 20,118.3 ± 11,138 μS/cm). The bicarbonate-rich groundwater is the coldest (25.6 ± 8 °C), but the chlorine-rich groundwater is the most acidic (pH = 6.6 ± 0.4; Table 1). The sulfate-rich groundwater has the lowest Cl^- (%), but the highest SO_4^{2-} (%) and temperature (Fig. 5b, c and h, respectively). The average arsenic, boron and electric conductivity values in this group (847.2 ± 679 μg/l, 14,678.8 ± 13,221 μg/l and 3,940.0 ± 540 μS/cm, respectively) lie between those of bicarbonate- and chlorine-rich groundwater. A part of this information cannot be obtained from categorizing the samples based on the dominant anion species (i.e. the right panels in Fig. 5a to 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, the chlorine-rich groundwater is not condensed in the corners of the ilr-ion plot and their chemical variation can be assessed in detail.

Excluding the sulfate-rich groundwater, the positive correlation between the samples is better demonstrated in Fig. 7b. The bicarbonate-rich groundwater shows 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 groundwater 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 the chlorine-rich groundwater whilst significant changes of $[Cl^- | SO_4^{2-}]$ is evident in the bicarbonate-rich groundwater (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 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. The Ca^{2+}/Mg^{2+} log-ratio indicates the smallest value in seawater and increases with decreasing Ca^{2+} and Mg^{2+} versus $Na^+ + K^+$ in chlorine-rich groundwater (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 the mixed groundwater, the bicarbonate-rich groundwater is enriched in bicarbonates. In accordance with seawater and thermal fluids, the chlorine-rich groundwater contains 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 the bicarbonate-rich groundwater samples are composed of equal chloride and sulfate contents. Their mixture with hydrothermal fluid and/or seawater results in increasing chloride, but steam-heating leads to elevated sulfate contents. 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, strength of the relationship between Na^+–Cl^-, B–Cl^- and B–Na^+ is evaluated via symmetric coordinates and log-transformed values (Fig. 8). In the groundwater samples 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 the bicarbonate- and chlorine-rich groundwater, respectively. The mixed groundwater represents a moderate R-squared value. It is noteworthy that the R^2 values are higher in Fig. 8b, d and f due to spurious correlation. For instance, approximately 80% ($p < 0.05$) of sodium variation is explained by chloride in the groundwater samples mainly derived from meteoric water ($C1_{BGW}$). It reduces to 45% ($p < 0.05$) when symmetric coordinates of Na^+ and Cl^- are utilized for the linear regression model which seems more reliable regarding the groundwater origin (Fig. 8a and b). Contrary to Fig. 8f, the B–Cl^- relationship in the chlorine-rich groundwater is not statistically significant in Fig. 8e which 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 concentrations in the Cl^--bearing high temperature springs probably explain 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 groundwater 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 the mixed groundwater. Seawater contribution is likely a process explaining the deviation from the mixing trend between the bicarbonate- and chlorine-rich groundwater (Fig. 9). Arsenic enrichment around Solfatara crater (the sulfate-rich groundwater spatially occurs in this area) is indicative of significant As in the vapor phase, condensing at shallow depth and giving rise to the steam-heated groundwater. It is in agreement with almost high As values in the fumarolic condensates of Solfatara (As~3,000 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- and sulfate-rich groundwater occurs in quadrant 3 of the graph ($[Na^+ | Cl^-] > 0$ and $[Na^+, Cl^- | HCO_3^-] < 0$), but the chlorine-rich groundwater lies in quadrant 1 ($[Na^+ | Cl^-] < 0$ and $[Na^+, Cl^- | HCO_3^-] > 0$). Nevertheless, the mixed groundwater occurs 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 groundwater. The chlorine-rich groundwater is 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).

Precursors of volcanic eruption are mainly identified via geochemical measurements regularly on the proximal emissions such as fumaroles, hot springs and volcanic lakes (Aiuppa et al., 2007; De Moor et al., 2016) or by geodetic and seismic monitoring (Patanè et al., 2003; Sparks, 2003; McNutt and Roman, 2015). However, the intermediate or remote hydrogeochemical monitoring of volcanoes are sometimes documented for identification the precursors (Armienta and De la Cruz-Reyna, 1995; Caracausi et al., 2003). Barbieri et al. (2021) recently indicated that the remote approach is likely easier and more reliable than sampling the proximal volcanic fluid emissions. They found an increase of B and Li concentrations before the 2014 Bárðarbunga eruption (~115 km from the monitoring station) and the 2018 M_w 5.0 earthquake (occurred ~80 km from the monitoring site) in central Iceland during the 2010–2018 groundwater monitoring. It indicates the ongoing investigation on finding better precursors of volcanic eruption through hydrogeochemical monitoring. Given the results of compositional data analysis in Figs. 7 to 10, the recent advances in statistical analysis deserve further attention to explore the full potential of groundwater hydrogeochemistry for delineating volcanic activity and the related seismic events. The new approaches might identify the most useful chemical variables that can validate the seismic and volcanic events over time, avoiding arbitrary selection of a hydrogeochemical precursor. This is of the utmost importance in active volcanic areas like Campi Flegrei, where densely populated areas are located inside and around craters.

5. 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 groundwater; (2) chlorine-rich groundwater; (3) sulfate-rich groundwater; and (4) mixed groundwater. The relationships between variables in each group is then explored using compositional data analysis and 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 indicates 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 groundwater. 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.

CRediT authorship contribution statement

Pooria Ebrahimi: Conceptualization, Software, Formal analysis, Investigation, Visualization, Writing - Original Draft, Writing - Review & Editing; **Annalise Guarino**: Writing - Review & Editing; **Vincenzo Allocca**: Writing - Review & Editing; **Stefano Caliro**: Investigation, Resources, Validation, Writing - Review & Editing; **Rosario Avino**: Investigation; **Emanuela Bagnato**: Investigation, Writing - Review & Editing; **Francesco Capecchiacci**: Investigation, Writing - Review & Editing; **Antonio Carandente**: Investigation; **Carmine Minopoli**: Investigation; **Alessandro Santi**: Investigation; **Stefano Albanese**: Resources, Supervision, Project administration, Writing - Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gexplo.2021.106922.

References

AGIP, 1987. Modello geotermico del sistemaflegreo. Sintesi, SEG-MESG Internal Report, 23 pp.

Aitchison, J., 1982. The statistical analysis of compositional data. J. R. Stat. Soc. Ser. B Methodol. 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). Chem. Geol. 229 (4), 313–330.

Aiuppa, A., Moretti, R., Federico, C., Giudice, G., Gurrieri, S., Liuzzo, M., Papale, P., Shinohara, H., Valenza, M., 2007. Forecasting Etna eruptions by real-time observation of volcanic gas composition. Geology 35, 1115–1118.

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). J. Geochem. Explor. 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.

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). Environ. Monit. Assess. 187 (12), 761.

Armienta, M.A., De la Cruz-Reyna, S., 1995. Some hydro-geochemical fluctuations observed in Mexico related to volcanic activity. Appl. Geochem. 10 (2), 215–227.

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. J. Volcanol. Geotherm. Res. 17 (1–4), 289–311.

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. J. Volcanol. Geotherm. Res. 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. Geochim. Cosmochim. Acta 52 (2), 475–483.

Barbieri, M., Franchini, S., Barberio, M.D., Billi, A., Boschetti, T., Giansante, L., Gori, F., Jónsson, S., Petitta, M., Skelton, A., Stockmann, G., 2021. Changes in groundwater trace element concentrations before seismic and volcanic activities in Iceland during 2010–2018. Sci. Total Environ. 793, 148635.

Bolognesi, L., Noto, P., Nuti, S., 1986. In: Studio chimico ed isotopico della Solfatara di Pozzuoli: ipotesi sull’origine e sulle temperature profonde dei fluidi, 41. Rendiconti della Societa Italiana di Mineralogia e Petrologia, pp. 281–295.

Buccianti, A., Grunsky, E., 2014. Compositional data analysis in geochemistry: are we sure to see what really occurs during natural processes? J. Geochem. Explor. 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). Geochim. Cosmochim. 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 Appl. Geophys. 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. J. Volcanol. Geotherm. Res. 76 (1–2), 63–82.

Caracausi, A., Italiano, F., Paonita, A., Rizzo, A., Nuccio, P.M., 2003. Evidence of deep magma degassing and ascent by geochemistry of peripheral gas emissions at Mount Etna (Italy): assessment of the magmatic reservoir pressure. J. Geophys. Res. Solid Earth 108 (B10).

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). Sci. Rep. 7 (1), 1–15.

Celico, P., De Gennaro, M., Pagano, D., Ronca, A., Stanzione, D., Vallario, A., 1987. In: 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, pp. 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. Boll. Soc. Geol. Ital. 111 (3–4), 409–422.

Celico, F., Esposito, L., Mancuso, M., 2001. Hydrodynamic and hydrochemical complexity of Naples urban area: some interpretations. Geol. Tec. Ambient. 2, 35–54.

Chandrasekharam, D., Bundschuh, J., 2002. Geochemistry of thermal waters and thermal gases. In: Chandrasekharam, D., Bundschuh, J. (Eds.), Geothermal Energy Resources for Developing Countries. Swets & Zeitlinger Publishers, Lisse, The Netherlands, pp. 253–267.

Chiodini, G., Frondini, F., Cardellini, C., Granieri, D., Marini, L., Ventura, G., 2001. CO2 degassing and energy release at Solfatara volcano, Campi Flegrei, Italy. J. Geophys. Res. 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). Geophys. Res. Lett. 30 (8).

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. J. Geophys. Res. 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.

Cortecci, G., Noto, P., Panichi, C., 1978. Environmental isotopic study of the Campi Flegrei (Naples, Italy) geothermal field. J. Hydrol. 36 (1–2), 143–159.

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. Sci. Rep. 5, 13100.

Dall’Aglio, M., Martini, M., Tonani, F., 1972. In: Rilevamento geochimico delle emanazioni vulcaniche dei Campi Flegrei, Quaderni De ‘La Ricerca Scientifica’, 83. CNR, pp. 152–181.

De Moor, J.M., Aiuppa, A., Pacheco, J., Avard, G., Kern, C., Liuzzo, M., Martínez, M., Giudice, G., Fischer, T.P., 2016. Short-period volcanic gas precursors to phreatic eruptions: insights from Poás Volcano, Costa Rica. Earth Planet. Sci. Lett. 442, 218–227.

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. J. Maps 14 (2), 230–241. https://doi.org/10.1080/17445647.2018.1454352.

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.

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. J. Volcanol. Geotherm. Res. 195 (1), 48–56.

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. Bull. Volcanol. 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. In: Goldschmidt2021• Virtual• 4-9 July.

Dominech, S., Yang, S., Aruta, A., Gramazio, A., Albanese, S., 2022. Multivariate analysis of dilution-corrected residuals to improve the interpretation of geochemical anomalies and determine their potential sources: the Mingardo River case study (Southern Italy). J. Geochem. Explor. 232, 106890.

Dotsika, E., 2015. H-O-C-S isotope and geochemical assessment of the geothermal area of Central Greece. J. Geochem. Explor. 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. Environ. Sci.: Water Res. Technol. 7 (6), 1060–1077.

Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Math. Geol. 37 (7), 795–828.

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

Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis. Springer International Publishing. https://doi.org/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. Rend. Soc. Ital. Mineral. Petrol. 43, 1139–1150.

Giggenbach, W.F., 1991. Chemical techniques in geothermal exploration (co-ordinator). In: D’Amore, F. (Ed.), Application of Geochemistry in Geothermal Reservoir Development. UNITAR, pp. 119–144.

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. Geochem. Geophys. Geosyst. 20 (11), 5544–5555.

Guglielminetti, M., 1986. Mofete geothermal field. Geothermics 15 (5–6), 781–790.

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. Math. Geosci. 49 (6), 777–796.

McNutt, S.R., Roman, D.C., 2015. Volcanic seismicity. In: The Encyclopedia of Volcanoes. Academic Press, pp. 1011–1034. https://doi.org/10.1016/B978-0-12-385938-9.00059-6.

McQuitty, L.L., 1966. Similarity analysis by reciprocal pairs for discrete and continuous data. Educ. Psychol. Meas. 26 (4), 825–831. https://doi.org/10.1177/001316446602600402.

Moeck, I.S., 2014. Catalog of geothermal play types based on geologic controls. Renew. Sust. Energ. Rev. 37, 867–882.

Patanè, D., De Gori, P., Chiarabba, C., Bonaccorso, A., 2003. Magma ascent and the pressurization of Mount Etna's volcanic system. Science 299, 2061–2063.

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. Proc. R. Soc. Lond. 60 (359-367), 489–498.

Peccerillo, A., 1985. Roman comagmatic province (central Italy): evidence for subduction-related magma genesis. Geology 13 (2), 103–106.

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. Int. J. Earth Sci. 103 (2), 401–421.

Piper, A.M., 1944. A graphic procedure in the geochemical interpretation of water-analyses. EOS Trans. Am. Geophys. Union 25 (6), 914–928.

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. Sci. Total Environ. 607, 965–971.

Rosi, M., Sbrana, A., 1987. Phlegraean Fields, Quaderni de "La ricerca scientifica", Consiglio Nazionale delle Ricerche, 9, 175 pp.

Scarpati, C., Perrotta, A., Lepore, S., Calvert, A., 2013. Eruptive history of Neapolitan volcanoes: constraints from 40Ar–39Ar dating. Geol. Mag. 150 (3), 412–425.

Sellerino, M., Forte, G., Ducci, D., 2019. Identification of the natural background levels in the Phlaegrean fields groundwater body (Southern Italy). J. Geochem. Explor. 200, 181–192.

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. J. Geochem. Explor. 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.

Sparks, R.S.J., 2003. Forecasting volcanic eruptions. Earth Planet. Sci. Lett. 210 (1–2), 1–15.

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. Geophys. Res. Lett. 42 (10), 3847–3854.

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. Chem. Geol. 194 (4), 245–274.

Valentino, G.M., Stanzione, D., 2004. Geochemical monitoring of the thermal waters of the Phlegraean Fields. J. Volcanol. Geotherm. Res. 133 (1–4), 261–289.

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. J. Volcanol. Geotherm. Res. 91 (2–4), 329–344.

Van den Boogaart, K.G., Tolosana-Delgado, R., 2013. Analyzing Compositional Data With R, Vol. 122. Springer, Berlin. https://doi.org/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. Int. J. Earth Sci. 103 (3), 801–819.

Washington, H.S., 1906. In: The Roman Comagmatic Region, Vol. 57. Carnegie Inst. of Washington, p. 199.

Publication No. 2

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. https://doi.org/10.1039/D0EW01123A.

Environmental
Science
Water Research & Technology

ROYAL SOCIETY
OF CHEMISTRY

PAPER

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d0ew01123a

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†

Pooria Ebrahimi,[a] Stefano Albanese, *[a] Leopoldo Esposito,[b]
Daniela Zuzolo[c] and Domenico Cicchella[c]

Received 29th December 2020,
Accepted 9th April 2021

DOI: 10.1039/d0ew01123a

rsc.li/es-water

Providing safe tap water has been a global concern. Water scarcity, the ever-increasing water demand, temporal variation of water consumption, aging urban water infrastructure and anthropogenic pressure on the water resources are the greatest challenges in effective water supply. In the present article, the waters exploited to be introduced in a water distribution system (*i.e.* input water) and tap waters are collected for determination of metal(loid)s, ions and physicochemical parameters. Seasonal variation is observed in the chemistry of the input waters. Further, the annual total dissolved solids (TDS) of the tap waters range from 200 to 1000 mg l^{-1} which stresses the importance of interconnections between urban water reservoirs for mixing different water types and adjusting water quality. It is complicated in populated cities like Naples with an old water distribution network, which also challenges setting up hydraulic and water quality models. The preliminary data visualization indicates the different chemical characteristics of some samples that are supposed to receive water from the same source. This might explain the difficulties in understanding the network layout in Naples. Thus, the compositional nature of chemical data was considered in hierarchical cluster analysis (HCA) to seasonally study water transfer between urban water reservoirs and define the source of tap water in each city area. The proposed method can preliminarily divide the pipe network into unique clusters and provide an overview of the relationships between different components when representative models cannot be set up due to limited information about network characteristics. Hence, advanced water distribution simulation and management is encouraged.

Water impact

This article suggests using tap water chemistry to identify the source of water in large old water distribution systems and accordingly divide the network into several clusters which enhance management practices. Depending on the available information, the advanced models can then be applied in the areas with the same water source.

1 Introduction

A water distribution system (WDS) is a network of water supply in which consumers, sources and tanks can be the vertices of the network while the connecting pipes, pumps and valves can be the edges.[1] The size of WDSs has increased over time due to urbanization. It can reach up to hundreds to thousands of nodes and links which substantially enhances the topological (*e.g.*, organization of nodes and links in different combinations) complexities. Further, WDSs are nonlinear dynamic systems and noticeable uncertainties exist in system behaviors, such as the stochasticity of the flow rate, pressure and water demand.[2]

Nowadays, computer programs (*e.g.*, EPANET[3] and EPANET-PMX[4]) can perform advanced hydraulic and water quality modelling for effective planning and controlling pipe

---

[a] *Department of Earth, Environmental and Resources Sciences, University of Naples Federico II, Napoli 80126, Italy. E-mail: stefano.albanese@unina.it*
[b] *Acqua Bene Comune Napoli Azienda Speciale, Napoli 80147, Italy*
[c] *Department of Science and Technology, University of Sannio, Benevento 82100, Italy*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0ew01123a

networks along with monitoring water quality within WDSs. Nonetheless, setting up such models is not quite simple because the characteristics of many components with a wide range of properties are needed and many water companies have a vague global vision towards their production and distribution.[5] Modern guidelines recommend dividing the pipe network into unique clusters called district metered areas (DMAs) before further assessment.[6] Several techniques have been introduced to achieve this aim. Bartolín *et al.*[7] constructed a minimum spanning tree to check the unconnected parts of a network, segmentation, broken pipes and conveyance capacity. Likewise, Tzatchkov[8] used depth-first and breadth-first based graph algorithms for identifying independent networks together with the errors in network data and applying WDS segmentation accordingly. In a generalized graph decomposition model, a network can be characterized by forests (tree structure) and cores (looped structure) as the two main elements which facilitate WDS analysis.[9] The network's tree structure was then separated from its looped core and the size of nonlinearity significantly reduced in the upgraded version of the model.[10] Sempewo *et al.*[11] and Di Nardo and Di Natale[12] presented the potential of spatial analysis for supporting metered area design in WDSs. Similarly, Izquierdo *et al.*[5] and Herrera *et al.*[13] considered graphical and vector information to design partitioning methods based on machine learning techniques. The clustering coefficient in graph theory can be an indicator of path redundancy in WDSs.[14] Perelman and Ostfeld[1] introduced topological clustering tools to divide a WDS regarding flow direction in pipes. Complex network clustering is another method for the same task.[15,16] The modularity-based approach[16] was modified by Giustolisi and Ridolfi[17] through applying a modularity index in multi-objective optimization to obtain various decomposition results for a WDS. Diao *et al.*[18] considered the hypothesis that WDSs have been expanded community-by-community during urban development, and detected the sub-networks that serve the corresponding communities.

DMAs were introduced in Italy at the end of the 20th century,[19] but they have not widely been used probably due to the complexity of DMA design operations and lack of the required information about the WDS.[20] Following many maintenance operations in 2005, Giugni *et al.*[20] revealed 67% water loss as a consequence of corroded steel pipes and high pressure patterns in the eastern sector of Naples (*i.e.* "Napoli Est"). Accordingly, they considered this part of the city as a DMA and performed hydraulic modelling of the water network using EPANET. The best performance was achieved in a scenario proposing six DMAs by placing 14 intercepting valves together with 9 pressure reducing valves. This solution can save approximately 34% of the water loss. Water consumption patterns were also investigated using different clustering techniques at NW Naples.[21]

Controlling the chemical characteristics of tap water is a common safety practice around the world. The companies in charge of managing the water networks annually carry out many chemical analyses with the aim of monitoring water quality and verifying its compatibility with national or local guidelines. The most advanced regulations usually determine the representative analytical parameters and the minimum number of measurements that must be carried out during a year. In Italy, a national legislative decree[22] is considered as a reference for monitoring tap water quality. Basically, routine and extended checks need to be performed by the companies distributing water. The former is based on determining a restricted number of parameters in specific checkpoints previously confirmed with the local health authority. The measurements are performed on a seasonal to a daily basis and the parameters include those reported by the European Union.[23] More stringent criteria are sometimes chosen for the precautionary principle. Furthermore, to ensure a higher level of health protection, the decree recommends monitoring additional parameters to better assess the characteristics of local reservoirs and the status of the hydro-drinking water distribution systems spanning the national territory. For example, vanadium and chlorites are monitored for sanitary reasons while residual disinfectant and hardness are measured as water quality indicators. Although huge analytical databases exist and constantly grow, they have not been used for any other purposes than "to protect human health from the adverse effects of any contamination of water intended for human consumption by ensuring that it is wholesome and clean" (as reported by the European Directive[6]). These data can be potentially used to study the variations in the amount of water proceeding from different sources. Moreover, the potential influence of construction materials of various infrastructures on the water chemistry can be investigated.

The clustering techniques discussed at the beginning of this section have introduced substantially useful insights towards WDS management. However, they have not been used to improve the efficiency of the old WDS in Naples, one of the most populated cities with a long history in Italy. The studies regarding WDS management are limited in the literature and they do not address the situation in the entire system. It might explain the complex WDS and lack of reliable information about the characteristics of the network. Therefore, the optimum operation of the WDS is questionable in the study area. The major aim of this research is to integrate multivariate statistics and compositional data analysis (CoDA) to better understand the WDS of Naples based on the geochemical fingerprints of tap waters. The results of this investigation are useful for partitioning the distribution system into unique clusters and obtaining a general overview about the WDS. It can be a decision support tool for developing hydraulic zones and alleviating the ongoing problems in water distribution operations.

2 Characteristics of the regional and urban water distribution systems

Acqua Bene Comune (ABC) Napoli is a company in charge of providing water services to over two million inhabitants in the city of Naples. The company supplies water to various third-party operators and municipalities; it serves about 280 000 delivery points to end users and supplies a daily average volume of 350 000 m^3 of water. The distribution system is based on more than 2000 km of distribution network.

The aqueduct of Naples was inaugurated on 10th May 1885. At the beginning, it was totally fed by the waters of the Serino Aqueduct proceeding from an area, 41 km east of the city, where the Acquaro-Pelosi and Urciuoli springs are located (Fig. 1a). The springs are fed by the carbonate reservoir of the Mt. Terminio massif[24] which belongs to the Apennine chain, a sequence of nappes (alternating siliciclastic and carbonate deposits) overthrusted towards N-NE. Specifically, the water is transported with a pressure line and a free surface covered channel (made of tuff, limestone and brick masonry) for about 60 km from the springs to a receiving tank at the top of the Cancello Hill. From the latter, pressure pipes depart to feed the urban water system and the Lufrano plant (checkpoint).

After the war period (1947–1948), exploitation of the Lufrano aquifer using a complex system of wells began to provide a reserve water supply (AIR) and meet the water demand in case of water scarcity or increasing drinking water consumption (Fig. 1a). The well field has been expanded until the 80s and the water reached the Capodimonte reservoir by a lifting station in Lufrano. The Lufrano aquifer is situated in the southern sector of the Campania Plain, a wide coastal area corresponding to a graben formed in the Upper Pliocene due to the sinking of mesozoic carbonates, and subsequently filled by quaternary alluvial, marine and pyroclastic deposits. This aquifer is fed by the water flowing from the carbonates of Mts. Avella in the east. Thus, Mg^{2+} and Ca^{2+} *vs.* Na^+ and K^+ tends to decrease in groundwater as it flows from the recharge area towards Lufrano.[25]

The Campano Aqueduct was also joined to the water network in 1958 to meet the increasing water demand in the eastern sector of the city (Fig. 1a). The water supplied by this aqueduct is exploited from the Biferno spring and from the Torano and Maretto springs on the northeastern side and Tyrrhenian side of the Mt. Matese carbonate massif in the Apennines chain, respectively. The well field of Ponte Tavano which is supplied by the carbonate reservoirs of Mts. Durazzano and located a few kilometers north of the Cancello Hill belongs to the latter aqueduct as well.

The Western Campania Aqueduct (ACO) has contributed to the water balance of the city supply system since 1998. At the slopes of a rocky limestone hill in the Lazio sub-Apennines, the water in ACO originates from the springs of the Gari river in Cassino (Latium). The water from Cassino is mixed with the waters proceeding from several well fields (*i.e.* Peccia, Monte Maggiore, S. Sofia, Monte Tifata and S. Prisco) together with the waters of the S. Bartolomeo and Sammucro springs and it is transported to the slopes of the S. Prisco Hill, northwest of the city of Caserta (Fig. 1a). The water that reaches the S. Prisco node is divided into four main large pipes that feed the aqueduct of Naples and the Campano Aqueduct storage tank of S. Clemente and Melito.

Between 2007 and 2009, the Cancello well field was also established by ARIN (the former ABC Napoli) in the municipality of S. Felice a Cancello to ensure water balance in the city (Fig. 1a).

Before reaching the water network in Naples, the input waters undergo several mixing processes (Fig. 2) which can change the geochemical fingerprint of the natural reservoirs. Moreover, the changes are favored by temporary water storage in a few urban reservoirs constructed from tuff, concrete or a combination of both. Specifically, in the year 2017, the city aqueduct counted on 7 interconnected urban reservoirs (Capodimonte, S. Sebastiano, Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli), with a total storage capacity of approximately 400 million liters, located at different altitudes ranging from sea level to 458 m.a.s.l. at Camaldoli hill.

The Scudillo and S. Giacomo urban reservoirs receive waters from the Serino Aqueduct. The S. Giacomo reservoir is also interconnected with the Capodimonte reservoir. The water in the S. Sebastiano reservoir originates from the Cancello well field, Campano Aqueduct and Ponte Tavano wells and, seasonally, from the Serino Aqueduct. The Capodimonte reservoir receives waters from ACO and AIR wells; however, Pianura, Cangiani and Camaldoli reservoirs are connected to the S. Giacomo reservoir (Fig. 2).

The Neapolitan water network has been built in different timespans depending on the expansion of the city and, consequently, it is possible to find pipelines made from very different materials such as gray cast iron, ductile iron, steel, and reinforced concrete. Most of the network developed before the 70s (the historical center and western sector of the city) is characterized by gray cast iron while bituminous steel is present in recently urbanized areas in the east of the city (Poggioreale, Secondigliano, Ponticelli and Barra). Frequent network failures in the latter areas are due to poor conservation.

3 Materials and methods

3.1 Sampling and analysis

The south central region of Italy that overlooks the Tyrrhenian Sea is characterized by dry summers and moderately rainy winters, especially along the coast. Rainfall increases significantly towards the inner parts (*i.e.* the Apennines chain) during the cold seasons and the annual average reaches up to 1800 mm.[26] Southern Italy, including the study area, experienced an overall reduction of precipitation in the last decades which resulted in decreasing water infiltration and groundwater recharge. In this article, the sampling campaigns were conducted in 2017 when winter (including December 2016) and spring were relatively dry. Summer was substantially hot and dry, except for September when the average rainfall was twice the climatological value. Rainfall intensity significantly varied in autumn. October was

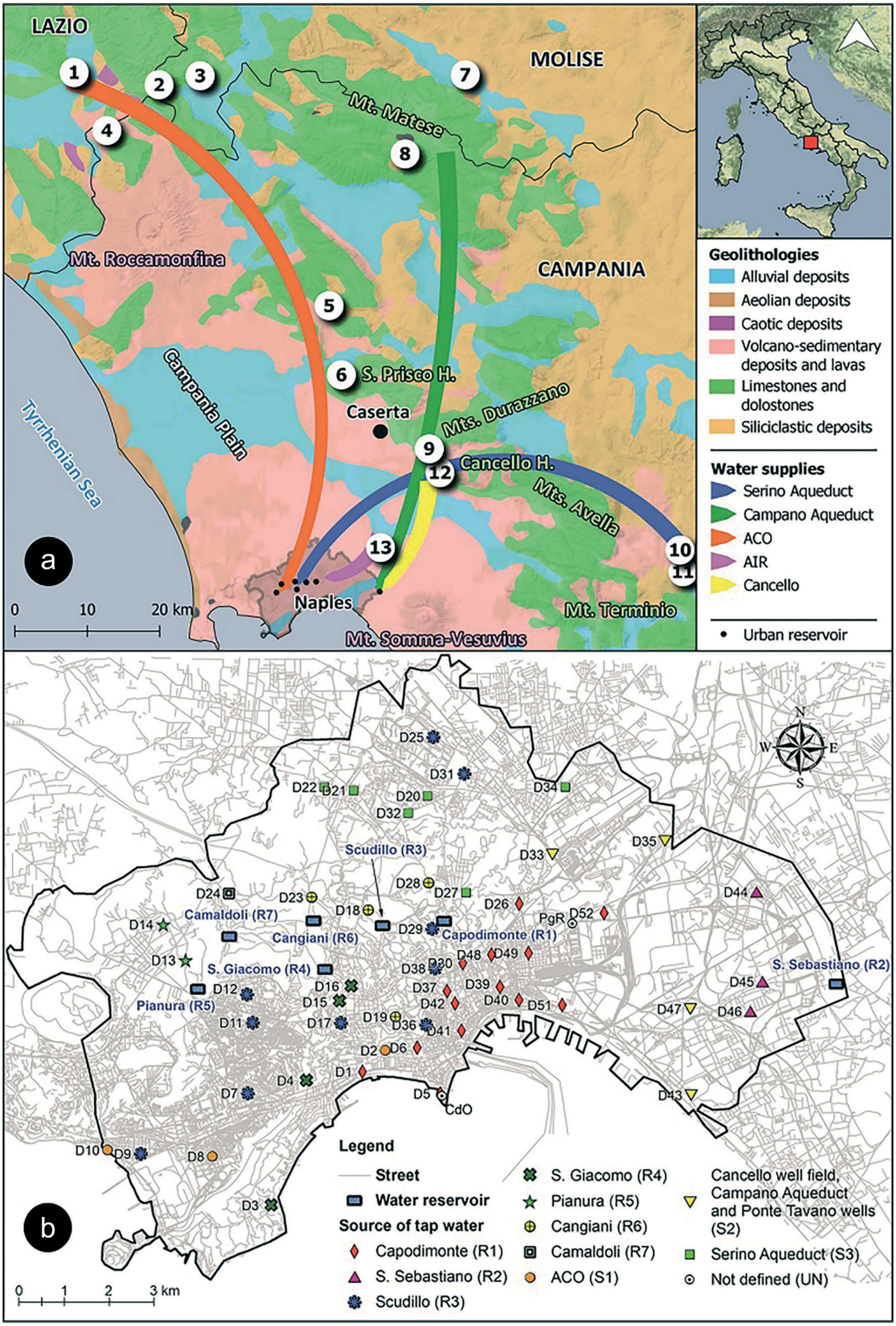


Fig. 1 (a) Lithology of the study area and the regional water supply to the Naples aqueduct. 1: Gari spring; 2: Sammucro spring; 3: S. Bartolomeo spring; 4: Peccia well field; 5: Monte Maggiore well field; 6: S. Prisco, S. Sofia, and Mt. Tifata well fields; 7: Biferno spring; 8: Torano and Maretto springs; 9: Ponte Tavano well field; 10: Urciuoli spring; 11: Acquaro-Pelosi springs; 12: Cancello well field; and 13: Lufrano well field. (b) Location of the collected tap waters (D1 to D52, CdO and PgR) and their source based on Fig. 2.

quite dry, but November was very rainy. Precipitation was in line with the expected climatological values in December.

To check water suitability for human consumption, NO_3^-, EC and pH were regularly measured (661 determinations for nitrate, but 664 measurements for EC and pH) in the input waters (I1 to I5 in Fig. 2) by qualified personnel of ABC Napoli. Moreover, concentrations of Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-}, Cl^- and F^- were determined monthly in AIR wells (I2) and seasonally in the other input waters. Further information about the number of samples is available in Table S2 and Fig. S1.† Tap water samples (Fig. 1b) were monthly collected from 53 sampling points (a total of 626 samples) across Naples after cleaning the bottles. Briefly, the bottles were rinsed several times with ultrapure water. The absence of both "memory" effect and release of chemicals from the container was then verified by filling randomly selected bottles with ultrapure water and checking the composition of the content analytically. The procedure has been approved by the Italian sole accreditation body (Accredia). In the sampling campaigns, the tap was turned on for at least 15 minutes (10 minutes more in the case of turbid water) and constant water temperature was attained prior to sampling. The bottles were first rinsed with tap water several times and then filled with water completely. The samples were stored at 5 ± 3 °C and transported to the Laboratorio Controllo Acque at ABC Napoli for chemical analyses. Ca^{2+}, Mg^{2+}, Na^+, K^+, NH_4^+, SO_4^{2-}, Cl^-, NO_2^-, NO_3^-, F^- and ClO_2^- were detected using ion chromatography (IC); however, HCO_3^-, Cl_2, TDS and turbidity were determined *via* titration, colorimetry, gravimetric method and nephelometry, respectively. Inductively coupled

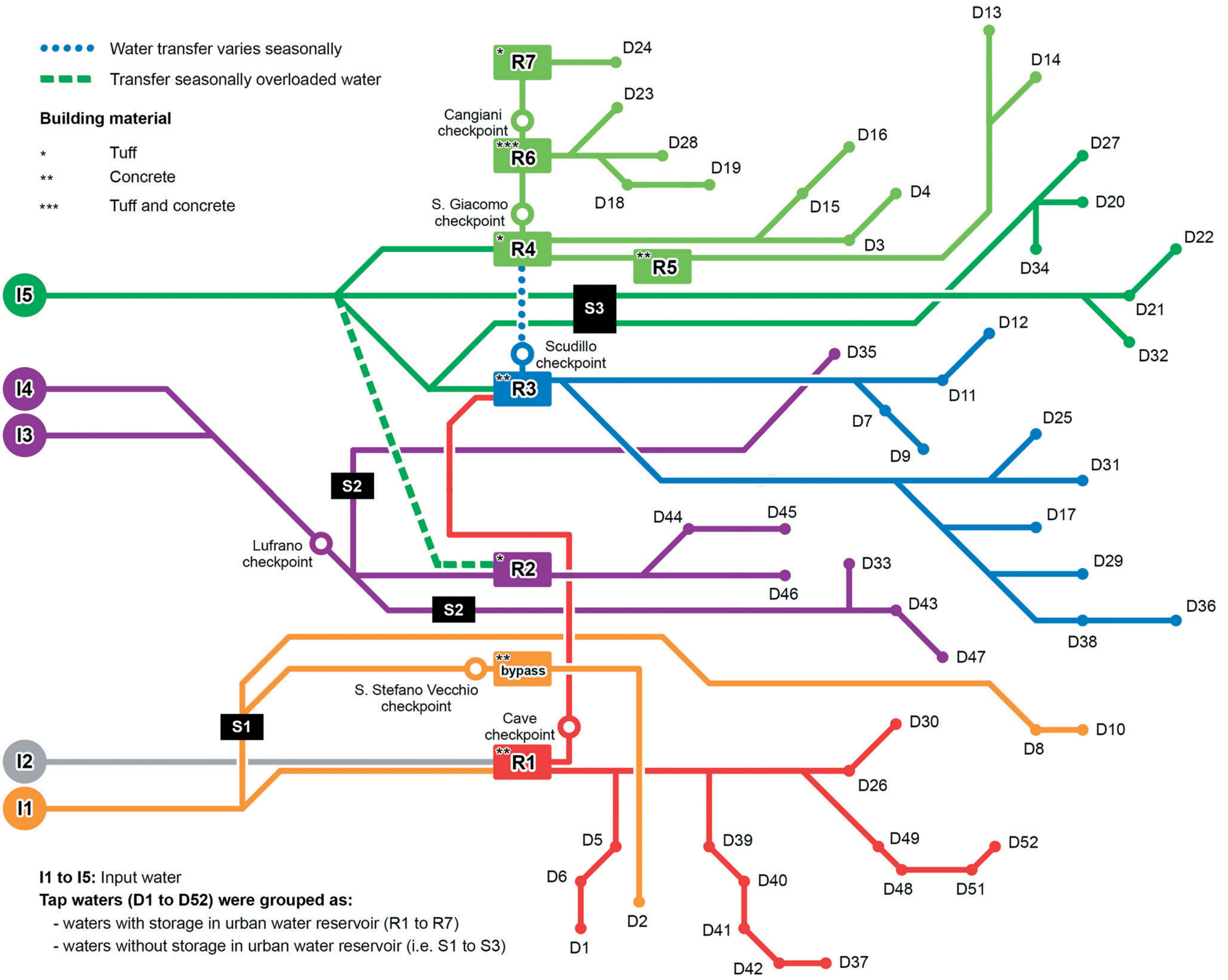


Fig. 2 The tap-water supply scheme of Naples. The input waters are defined as I1: ACO; I2: AIR wells; I3: Cancello well field; I4: Campano Aqueduct and Ponte Tavano wells; and I5: Serino Aqueduct. D1 to D52 represent the tap water samples. Every month, the local water company collects tap water at these stations to check tap water quality and the results are accessible to the public. The coordinates and address of these stations are provided in Table S1.† In this study, the tap waters stored in urban water reservoirs are grouped as R1 to R7 and the rest are categorized as S1 to S3. The urban water reservoirs are shown as R1: Capodimonte (92 m.a.s.l.); R2: S. Sebastiano (112 m.a.s.l.); R3: Scudillo (183 m. a.s.l.); R4: S. Giacomo (230 m.a.s.l.); R5: Pianura (220 m.a.s.l.); R6: Cangiani (303 m.a.s.l.); and R7: Camaldoli (458 m.a.s.l.). However, S1, S2 and S3 are defined as follows: S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; and S3: Serino Aqueduct.

plasma-mass spectrometry (ICP-MS) was also used for measurement of metal(loid)s (As, Cu, Pb, Cd, Ni, Cr, V, Hg, Fe, Mn and Al). Determination of electric conductivity (EC) and pH with a Mettler Toledo T70, and temperature with a digital thermometer was done in the field. The following formula which accounts for Ca^{2+} and Mg^{2+} as the main factors for total hardness (TH) was also applied for TH computation:[27]

$$\mathrm{TH} = 2.5\mathrm{Ca}^{2+} + 4.1\mathrm{Mg}^{2+} \tag{1}$$

where all values were in mg l^{-1}.

The detection limits were calculated as three times the standard deviation of the analyte concentration in the samples with very low content of the analyte: 0.01 μg l^{-1} for Hg, 0.1 μg l^{-1} for As, Pb, Cd, Ni, Cr, V, Mn and Al, 1 μg l^{-1} for Cu, Fe and ClO_2^-, 0.01 mg l^{-1} for NO_2^-, 0.02 mg l^{-1} for NH_4^+ and Cl_2, 0.03 mg l^{-1} for F^-, 0.07 mg l^{-1} for NO_3^-, 0.1 mg l^{-1} for Ca^{2+}, Mg^{2+} and SO_4^{2-}, 0.2 mg l^{-1} for Na^+, 0.25 mg l^{-1} for Cl^-, 0.4 mg l^{-1} for K^+, 2 mg l^{-1} for HCO_3^-, 10 mg l^{-1} for TDS, 1 μS cm^{-1} for EC and 0.05 NTU for turbidity. Blanks, trace element matrix reference materials (RMs) and NIST standard reference materials (SRMs) were employed to verify measurement precision and accuracy. The maximum calculated error is 11% and relative standard deviations (% RSDs) of the measurements range between 0 and 7%.

3.2 Data preparation and statistical analysis

Different steps from data treatment to data visualization and statistical analysis are summarized in Fig. 3 and further discussed in this section. Once the experimental results were available, the samples with unacceptable cation–anion balance (CAB) error (*i.e.* not within ±10%) were omitted and

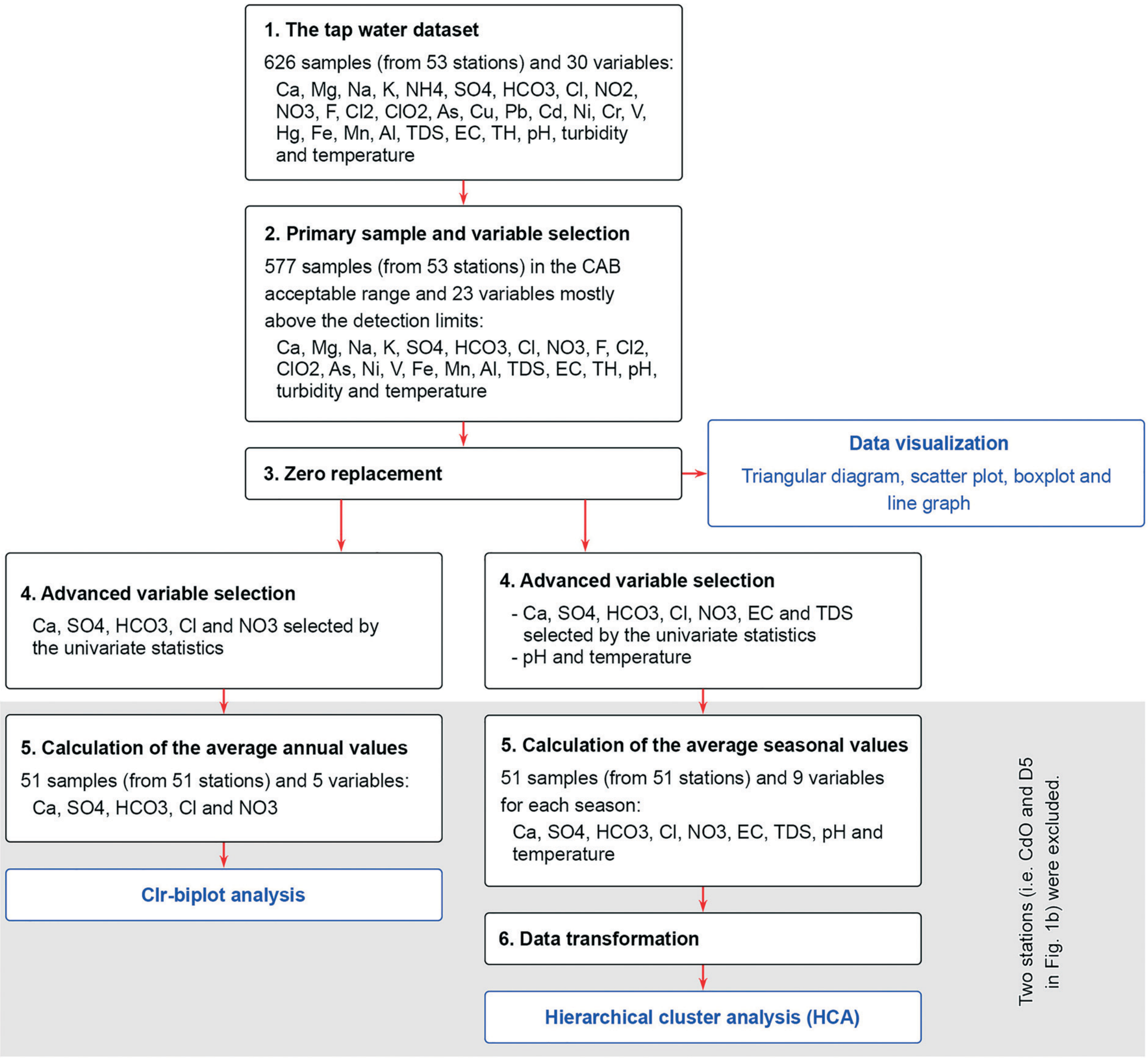


Fig. 3 The flowchart depicting different steps of data treatment before data visualization and statistical analysis.

577 tap water samples remained. Cu, Pb, Hg, Cd, Cr, NO_2^- and NH_4^+ were ignored in this investigation because most of the values were below the corresponding detection limits (Fig. 3).

Combined with conventional statistical analyses, compositional data analysis is a powerful tool to get better knowledge of the relationships and dependencies between elements, various geochemical processes and different sources of chemicals.[28,29] Thus, the hierarchical clustering algorithm was coupled with CoDA in this article. To apply compositional data analysis and statistical analysis, a dataset with no value below the detection limits is necessary. Because the proportion of rounded zeros (the values below the detection limit) for Ni, Fe, Mn, Al and ClO_2^- is higher than 10%, a complete data matrix with nonzero entries was obtained through running a multivariate modified EM algorithm after additive log-ratio transformation of the compositions (Fig. 3).[30] This parametric approach consists of modifying the EM-algorithm in order to include the (logratio) imputed values with smaller values than the (logratio) detection limit. The zCompositions package of R software was utilized for zero replacement.

The chemicals that feasibly induce adverse effects on the statistical analysis were then culled using univariate statistics (Table 1). To be very concerned with the deleterious effects of measurement errors, the problematic group of chemicals were detected using the proportion of concentrations less than three times the first instrumental calibration point. It was assumed that such low concentrations have noticeable errors. After culling the analytes exceeding the maximum acceptable limit of 10%, Ca^{2+}, SO_4^{2-}, HCO_3^-, Cl^-, NO_3^-, EC and TDS were retained. In addition, pH and temperature that have no detection limits were included. TH was also ignored because of using Mg^{2+} with approximately 19% of values below three times the first instrumental calibration point to calculate the parameter (Table 1; Fig. 3).

Because it is expected that precipitation and evaporation influence the chemistry of input waters, the seasonal geometric means of the monthly determined variables were subsequently calculated for each station (Fig. 3) to develop this study based on each season. Two sampling stations (*i.e.* CdO and D5) without analytical results in some seasons were also ignored to keep the same number of samples between seasons. Due to the compositional nature of chemical data, the ratio between chemicals provides more information than the absolute measured values.[31] Therefore, the dataset was transformed by means of the centered log-ratio (clr) transformation introduced by Aitchison.[32] The advantage of clr- over alr- (additive logratio[32]) transformation is that it represents each variable as a ratio to a central value (the geometric mean of a subcomposition) instead of favoring one variable (Dth part) in all denominators.[33] However, this approach can lead to some difficulties in interpretation of clr values as the concentration of the original elements because any single component can be influenced by dilution, enrichment and depletion in the set of components considered. Nevertheless, clr-transformation was chosen because it is effective for understanding the underlying geochemical processes.[34] Specifically, the clr-transformation for a composition $x = [x_1, x_2, \dots, x_D] \in S^D$ is the transformed data $y \in R^D$ with:[35]

Table 1 Statistical summary of the measured parameters in the selected 577 tap water samples

Parameter	FCP[a]	% < 3FCP	Min	5% ile	25% ile	50% ile[b]	75% ile	95% ile	Max	IQR	Average
As (μg l^{-1})	1.0	42.1	<D.L.[c]	1.0	2.0	3.0	3.0	3.0	4.0	1.0	2.4
Ni (μg l^{-1})	1.0	86.8	<D.L.	0.4	0.6	1.0	1.0	5.0	13.0	0.4	1.4
V (μg l^{-1})	1.0	23.2	1.0	2.0	3.0	3.0	4.0	6.0	8.0	1.0	3.5
Fe (μg l^{-1})	1.0	37.6	<D.L.	0.6	2.0	4.0	6.0	18.0	122.0	4.0	5.9
Mn (μg l^{-1})	1.0	58.1	<D.L.	0.5	1.0	2.0	4.0	9.2	29.0	3.0	3.1
Al (μg l^{-1})	1.0	85.1	<D.L.	0.3	0.4	1.0	2.0	5.0	140.0	1.6	2.0
ClO_2^- (μg l^{-1})	50.0	94.3	<D.L.	1.6	4.1	6.6	71.0	152.6	312.0	66.9	39.3
Cl_2 (mg l^{-1})	0.08	84.2	<D.L.	0.11	0.16	0.2	0.22	0.26	0.44	0.06	0.2
Ca^{2+} (mg l^{-1})	12.0	0.0	46.4	56.4	100.9	121.3	132.9	180.3	204.1	32.0	116.5
Mg^{2+} (mg l^{-1})	6.0	18.5	6.9	9.6	24.1	30.1	32.1	45.0	49.9	8.0	27.8
Na^+ (mg l^{-1})	6.0	36.7	4.0	7.0	13.0	23.0	36.0	93.0	108.0	23.0	29.3
K^+ (mg l^{-1})	3.0	36.2	1.0	3.0	6.3	11.6	14.4	18.9	23.9	8.1	10.9
SO_4^{2-} (mg l^{-1})	1.0	0.7	1.0	9.0	19.0	36.0	46.0	55.2	64.0	27.0	33.1
HCO_3^- (mg l^{-1})	2.0	0.0	183.0	221.6	411.0	505.0	534.0	698.0	754.0	123.0	472.6
Cl^- (mg l^{-1})	0.5	0.0	8.0	9.0	16.0	31.0	53.0	184.0	219.0	37.0	47.4
F^- (mg l^{-1})	0.09	27.9	0.1	0.2	0.2	0.5	0.6	0.7	0.9	0.3	0.4
NO_3^- (mg l^{-1})	1.0	0.3	<D.L.	4.0	8.0	17.0	22.0	32.0	38.0	14.0	16.5
EC (μS cm^{-1})	1.0	0.0	259.0	345.0	651.0	798.0	916.0	1405.6	1610.0	265.0	789.5
TDS (mg l^{-1})	10.0	0.0	181.0	242.0	456.0	559.0	641.0	984.4	1127.0	185.0	552.7
TH (mg l^{-1})	—	—	149.8	180.5	361.6	425.0	462.9	632.0	713.6	101.3	405.1
Turbidity (NTU)	0.05	24.6	<D.L.	0.1	0.2	0.2	0.3	0.4	1.1	0.1	0.2
pH	—	—	6.6	6.8	7.1	7.3	7.5	7.9	8.4	0.4	7.3
T (°C)	—	—	11.1	11.9	13.7	14.4	15.0	17.0	23.5	1.3	14.4

[a] The first instrumental calibration point. [b] Median. [c] < Detection limit.

$$y = \mathrm{clr}(x) = \left(\ln \frac{x_1}{g(x)}, \ln \frac{x_2}{g(x)}, \ \ldots, \ln \frac{x_D}{g(x)} \right) = \ln \frac{x}{g(x)} \quad (2)$$

where $g(x)$ is the geometric mean across the composition x which can be computed as:

$$g(x) = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} = \sqrt[n]{x_1 \cdot x_2 \ldots x_n} = \exp\left[\frac{1}{n} \sum_{i=1}^{n} \ln x_i \right] \quad (3)$$

EC, TDS and temperature that cannot be treated like compositional data were log-transformed and no transformation was applied to pH due to its logarithmic scale (Fig. 3). Finally, for hierarchical cluster analysis (HCA), the Euclidean distance (for computation of the distances between the samples) and Ward's method (for agglomerative clustering) were used as the most widely applied methods that form easily interpretable and possibly significant clusters in the hydrogeochemical, hydrologic and geologic contexts (Fig. 3).[36–38] The CoDaPack v2.02.21 and "cluster" R package were used for the data transformation and HCA, respectively.

In order to better understand the compositional variability, the associations between the parts of a (sub)composition and the likely natural sources of tap water, a clr-biplot was also generated using the annual geometric mean of Ca^{2+}, HCO_3^-, Cl^-, SO_4^{2-} and NO_3^- in each station *via* the CoDaPack v2.02.21. Biplots graphically represent the clr-transformed compositional data based on the principal component analysis.[39] Construction of a clr-biplot starts with clr transformation of the data matrix *via* eqn (2). The clr-components are then centered in order to put the origin of axes in the center of the data. In the next step, a singular value decomposition (SVD) of the centered clr-data is conducted to generate orthogonal axes showing maximum variability of the dataset. The observations and the variables are finally demonstrated as dots and rays (vectors) on a plot, respectively.[40] The data points with similar values will be clustered on biplots. The squared length of a clr-ray is comparable with the clr-variance of the corresponding variable. Nevertheless, the squared length of the link between two vertices is almost equal to the variance of the log-ratio of those two parts. Hence, the correlation of two variables increases with decreasing distance between the corresponding vertices.[41,42] Uncorrelated components are at right angles to each other. Further, the cosine of the angle between two intersecting links explains the correlation coefficient between the log-ratios. Consequently, if two links are at right angles, zero correlation of the log-ratios is expected.[31]

4 Results and discussion

4.1 Chemistry of the input water

According to Fig. 2, the water from ACO (I1) and AIR wells (I2) is delivered to the Capodimonte reservoir (R1). Ca^{2+}, Na^+, Cl^-, SO_4^{2-}, NO_3^-, F^- and EC values are low in the former and high in the latter which result in their different chemical patterns (Table S2 and Fig. 4, S1 and S2†). The highest concentrations of SO_4^{2-}, NO_3^- and F^- are determined in I2; however, I3 (Cancello well field) and I4 (Campano Aqueduct and Ponte Tavano wells) account for the highest values of the other parameters (Table S2† and Fig. 4). The analytical results of I3 and I4 are almost the same leading to their similar chemical signature in Fig. 4. The lowest chemical contents are reported for the slightly basic waters originating from the Serino Aqueduct (I5). The pH values indicate the neutral characteristics of the other input waters. Statistical distributions of nitrate, EC and pH which were measured regularly over the timespan of this study indicate that composition of the input water probably reflects precipitation and evaporation in the study area (Fig. S1†). Mixing different proportions of the waters with high (*i.e.* I2, I3 and I4) and low (*i.e.* I1 and I5) dissolved contents plays an important role in meeting the quality standards.[22] However, changes in volume and composition of the input water during the year can be a great challenge in water transfer between the urban water reservoirs to satisfy the increasing water demand of the inhabitants (Table S2 and Fig. S1 and S2†).

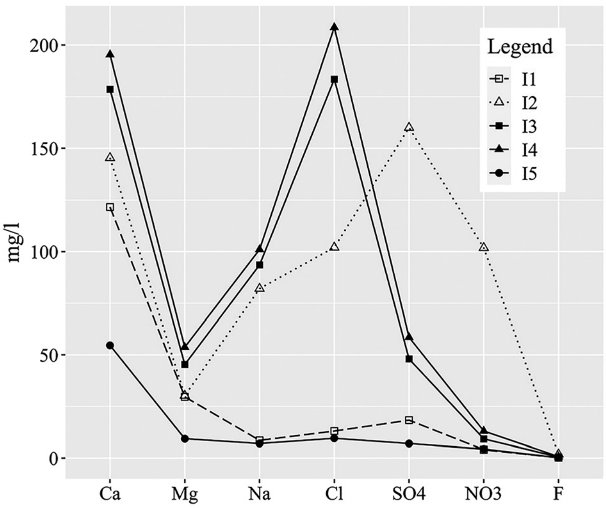


Fig. 4 Chemical compositions of the input waters in the water distribution system of Naples in 2017. I1: ACO; I2: AIR wells; I3: Cancello well field; I4: Campano Aqueduct and Ponte Tavano wells; and I5: Serino Aqueduct.

4.2 Chemistry of tap water

The chemical characteristics of the collected tap waters are statistically summarized in Table 1. Even though the average concentrations of cations are in the decreasing order of Ca^{2+} (116.5 mg l^{-1}) > Na^+ (29.3 mg l^{-1}) > Mg^{2+} (27.8 mg l^{-1}) > K^+ (10.9 mg l^{-1}), the mean contents of anions follow the order of HCO_3^- (472.6 mg l^{-1}) > Cl^- (47.4 mg l^{-1}) > SO_4^{2-} (33.1 mg l^{-1}) > NO_3^- (16.5 mg l^{-1}) > F^- (0.4 mg l^{-1}). On average, the values of the selected metal(loid)s (As, Ni, V, Fe, Mn and Al) are below 6 μg l^{-1}. Chlorine dioxide (ClO_2^-) and chlorine (Cl_2), which are utilized for chemical disinfection of water, do not exceed 153 μg l^{-1} and 0.3 mg l^{-1}, respectively, in 95% of the samples. The average and median of Cl_2, Ca^{2+}, Mg^{2+}, K^+,

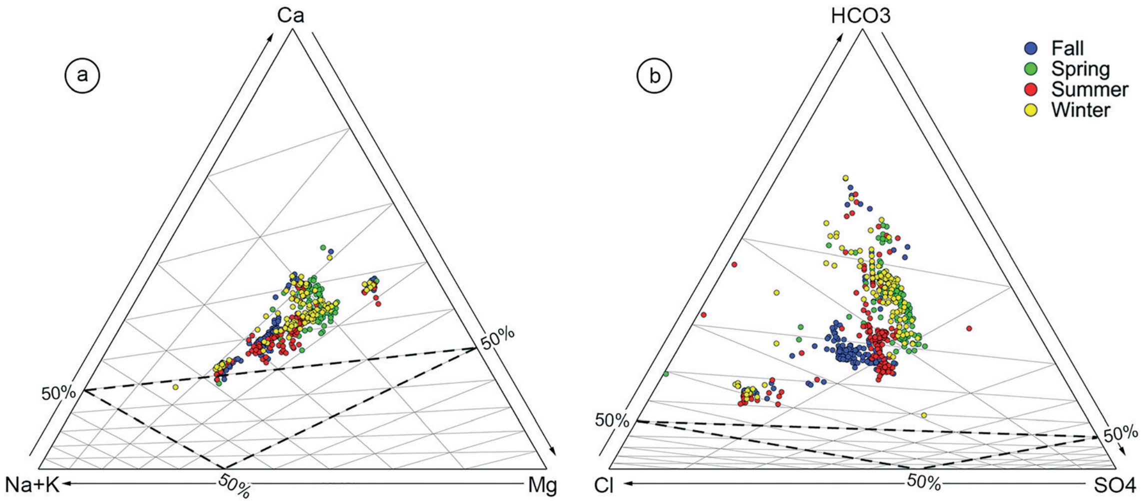


Fig. 5 Diagrams indicating seasonal composition of tap waters: (a) relative abundance of major cations; (b) relative abundance of major anions. The data are centered to get a better view of the pattern.

SO_4^{2-}, HCO_3^-, NO_3^-, EC, TDS, TH, pH and temperature are approximately equal. Tap waters are colder than 23.5 °C and indicate slightly acidic to basic nature (pH = 6.6 to 8.4). The interquartile range (IQR) explains the variance of the measured parameters. In this article, metal(loid)s, Cl_2, F^-, turbidity, pH and temperature have the lowest variation (IQR < 5). Whilst most major ions indicate medium variance (IQR = 5–100), the IQR values of EC, TDS, TH and bicarbonate are the greatest (>100) (Table 1).

Concentrations of major cations (*i.e.* Ca^{2+}, Mg^{2+}, Na^+ and K^+) and anions (*i.e.* Cl^-, HCO_3^- and SO_4^{2-}) were plotted on trilinear diagrams (Fig. 5) to evaluate seasonal variation of tap water chemistry based on the dominant cations and anions. Fig. 5a and b indicate that Ca^{2+} and HCO_3^- are the dominant cation and anion, respectively. In summer and fall, the relative concentration of $Na^+ + K^+$ and Cl^+ increases. Comparison of Ca^{2+}, Mg^{2+}, Na^+, SO_4^{2-}, Cl^-, NO_3^- and F^- measurements in the input waters (I1 to I5 in Fig. S2†) and the tap waters (R1 to R7 and S1 to S3 in Fig. S3†) help to interpret the seasonal changes in water chemistry. Except for S. Sebastiano (R2), ACO (S1) and Serino Aqueduct (S3) waters, this seasonal trend is confirmed in the rest of the collected tap waters (Fig. S3†). Considerable Na^+ and Cl^+ contents in the water from AIR wells (I2; Fig. S2b†) might explain the high proportion of I2 in Capodimonte (R1) tap waters in summer and fall. This chemical signature can subsequently affect the relative abundance of sodium and chloride in the other waters due to the interconnections of the water reservoirs. Summer is the dry season in the study area with less precipitation and more groundwater exploitation because of high water demand. Therefore, the geochemical characteristics of pyroclastic materials in the Lufrano aquifer might significantly influence Na^+ and Cl^- contents in the water from AIR wells (I2).[25] Excess Na^+, K^+ and Cl^- in fall can be explained by the proportion of I2 in R1 along with the time lag between precipitation and groundwater recharge.

4.3 Construction materials of the reservoirs and tap water chemistry

Plotting the geometric mean of TDS against TH demonstrates that although the highest values are observed in the S. Sebastiano reservoir (R2) and Cancello well field, Campano Aqueduct and Ponte Tavano wells (S2), the lowest values are

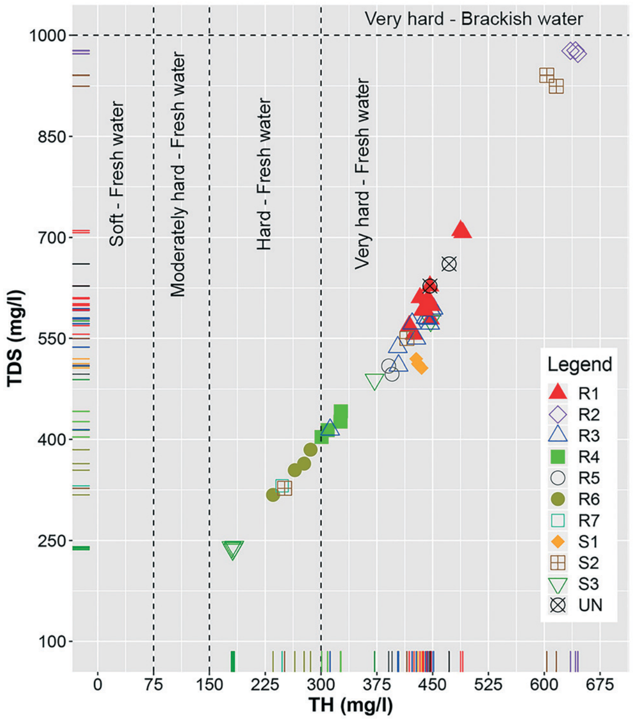


Fig. 6 Scatter plot of TDS *vs.* TH in the tap waters. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

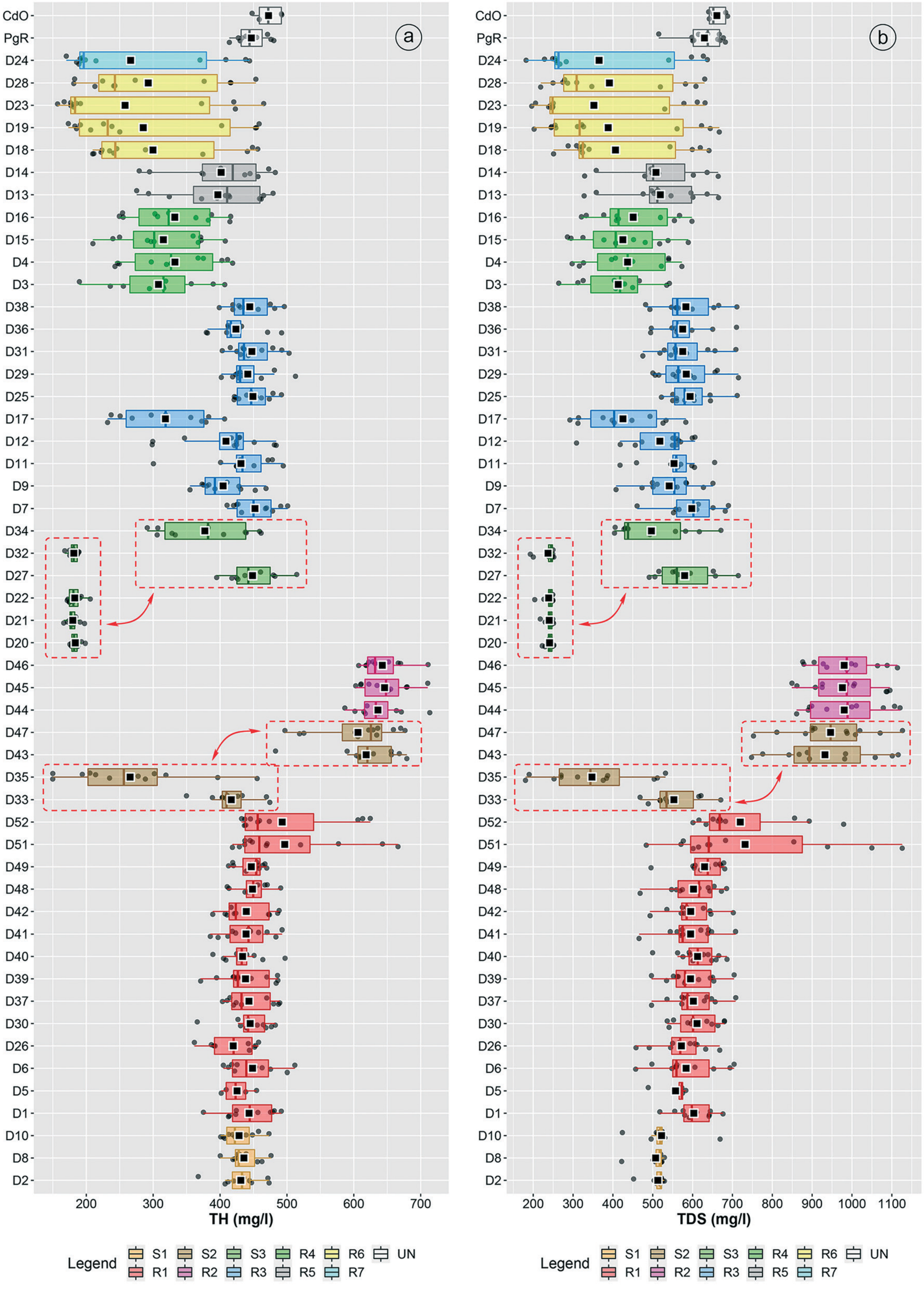


Fig. 7 Variation of TH (a) and TDS (b) in the collected tap water samples. The gray circles indicate the monthly measurements and the black squares show the annual average values. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

determined in the Serino Aqueduct (Fig. 6). The composition of the tap waters from Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli reservoirs (R3, R4, R5, R6 and R7, respectively) ranges between the chemical characteristics of the tap waters from the Serino Aqueduct and Capodimonte reservoir (Fig. 6). It suggests a proportional negative influence of the water from Capodimonte on the tap water quality because the reservoirs are extensively interconnected. In a decreasing order, Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli tap waters are feasibly influenced by Capodimonte water depending on the transfer rate of water from the Capodimonte reservoir to the Camaldoli reservoir (Fig. 6). Thus, there is more than one unknown factor restricting the ability to thoroughly evaluate the effect of building materials of the reservoirs on the water chemistry. Regardless of water storage in the S. Sebastiano reservoir, most of the R2 and S2 tap waters distributed in the eastern part of the city reveal elevated TH and TDS values (Fig. 7). The urban water distribution network (Fig. 2) demonstrates that the tap waters originating from the S. Sebastiano reservoir (R2) are occasionally mixed with the water from the Serino Aqueduct (I5) preventing a comprehensive interpretation of the influence of tuff on water quality. Fig. 6 also shows that the collected tap waters are hard-fresh water and very hard-fresh water according to the TDS classification of Todd[27] and the TH classification of Sawyer *et al.*[43]

Because the Capodimonte reservoir does not receive water from many sources (Fig. 2), high values of anions, cations and the other parameters in R1 can be contributed by the input waters (I1 and I2) or the chemical interactions of the water with the concrete reservoir (Fig. 6, 7 and S4†). Compared with ACO (I1), elevated concentrations of the measured ions in AIR wells (I2) can be considered as the main source of chemicals in the Capodimonte tap waters (Fig. S2a and b†). ACO (S1), which is not mixed with AIR wells (I2), has low contents of ions. A Portland cement plaster was applied to the interior tuffaceous walls of the Capodimonte reservoir. The hydrated phases of cement paste are not stable at pH values below 12 to 13. Prolonged contact of the porous material with near-neutral waters leads to dissolution and leaching of the soluble components.[44] The highly alkaline pore solution of a typical Portland cement paste leads to the beginning of the leaching process by removing Na^+ and K^+. It is followed by dissolution of portlandite and subsequent leaching of calcium from silicates.[45–47] Further, the dissolution/precipitation processes of aluminate phases including AFm (Al_2O_3–Fe_2O_3-mono: family of lamellar calcium alumina–ferrite hydrates), ettringite and calcite are affected.[48,49] Na^+ and K^+ contents are higher in the water supplied by the Capodimonte reservoir (R1) than those of ACO (S1) probably suggesting the first stage of chemical reactions between the water and the concrete reservoir (Fig. S4i and j†). It is in accordance with the indifferent concentrations of Ca^{2+} and Al in these two tap water sources (Fig. S4f and k†). However, further investigation is needed to confirm this assumption. The reported leached depth of 5 to 10 mm in concretes submerged in still natural water for 100 years may explain the slow rate of the leaching process.[50] Thus, elevated concentrations of Cl^-, SO_4^{2-}, NO_3^-, and F^- in the water provided by Capodimonte highlight the unique role that AIR wells play in the water chemistry (Fig. 2, S3 and S4†).

According to Fig. 7, TH and TDS values of D33 and D35 are strikingly different from those of D43 and D47. Likewise, variation of TH and TDS in D27 and D34 is completely different from D20, D21, D22 and D32 supplied by the same source (Fig. 2). Similar surprising patterns in the other parameters (Fig. S4†) might indicate lack of information about the network structure in some parts of Naples and stress the need for characterizing the source of tap water through the existing dataset.

4.4 Chemical fingerprint of the source in the water distribution network

A clr-biplot was used (Fig. 8) to investigate the compositional variability along with the associations between the parts of a (sub)composition and to identify the probable natural sources of the water provided for human consumption. The first two principal components explain 98% (62 and 36%, respectively) of the total cumulative variability being an acceptable approximation for the 5-dimensional scatterplot of the clr-transformed data. Chloride and nitrate indicate the greatest compositional variability; however, sulfate shows the smallest variability. Ca^{2+} and HCO_3^- are significantly

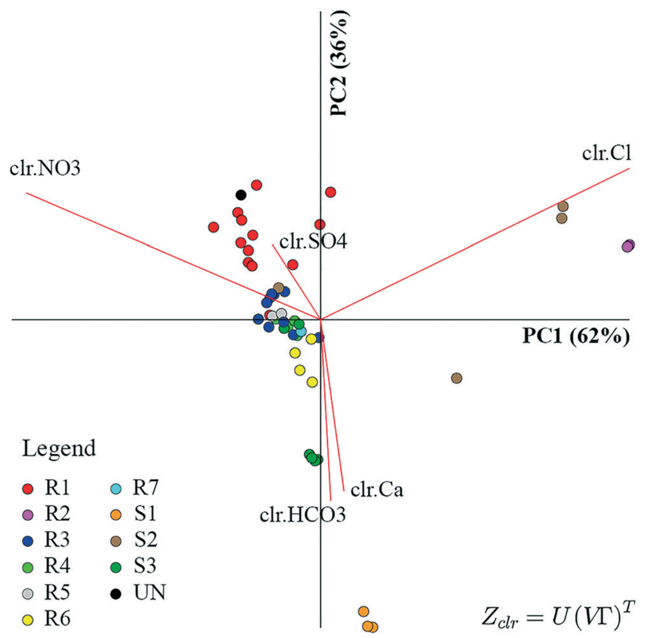


Fig. 8 The clr-biplot of the selected ions retaining 98% of the total variability. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

correlated, but their correlation with SO_4^{2-} is negligible. The clr-biplot indicates that the tap water samples can be discriminated by three groups of variables. The samples from the S. Sebastiano reservoir (R2) together with the Cancello well field, Campano Aqueduct and Ponte Tavano wells (S2) are plotted around the "clr.Cl" ray introducing chloride as the main chemical fingerprint of the waters. Sulfate and nitrate characterize the hydrochemistry of Capodimonte water (R1) whilst calcium and bicarbonate are the main characteristics of S1 and S3 waters (ACO and Serino Aqueduct, respectively) (Fig. 8). Occurrence of the other samples (Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli) between these two end-members confirms the results in section 4.3 and Fig. 6.

Visualizing the spatial distribution of the factor scores of the principal components, after appropriate classification, also depicts the chemical end-members (Fig. 9). Cl^- and NO_3^- are the ions with significant factor loadings in the first principal component (PC1) (0.718 and −0.684, respectively). According to Fig. 9a, the highest PC1 factor scores in the eastern part of the city correspond to high chloride content in R2 and S2 tap waters. Nonetheless, nitrate enrichment in most of the water samples from Capodimonte, Scudillo, Pianura, S. Giacomo, Cangiani and Camaldoli is reflected in their lowest factor scores. In terms of Cl^- and NO_3^-, the S1 and S3 tap waters (Serino Aqueduct and ACO, respectively) lie between the above-mentioned end-members (Fig. 9a).

On the other hand, the second principal component (PC2) accounts for significant negative factor loadings of Ca^{2+} and HCO_3^- (−0.525 and −0.555, respectively), and a high positive factor loading of Cl^- (0.463). The group with the highest factor scores is representative of the Cl-bearing R1, R2 and S2 tap waters. However, the Serino Aqueduct (S3) and ACO (S1) are characterized with the lowest factor scores due to the unique clr-transformed calcium and bicarbonate values. The rest of the samples feasibly depict a mixed source (Fig. 9b).

4.5 Hierarchical cluster analysis and tap-water supply map

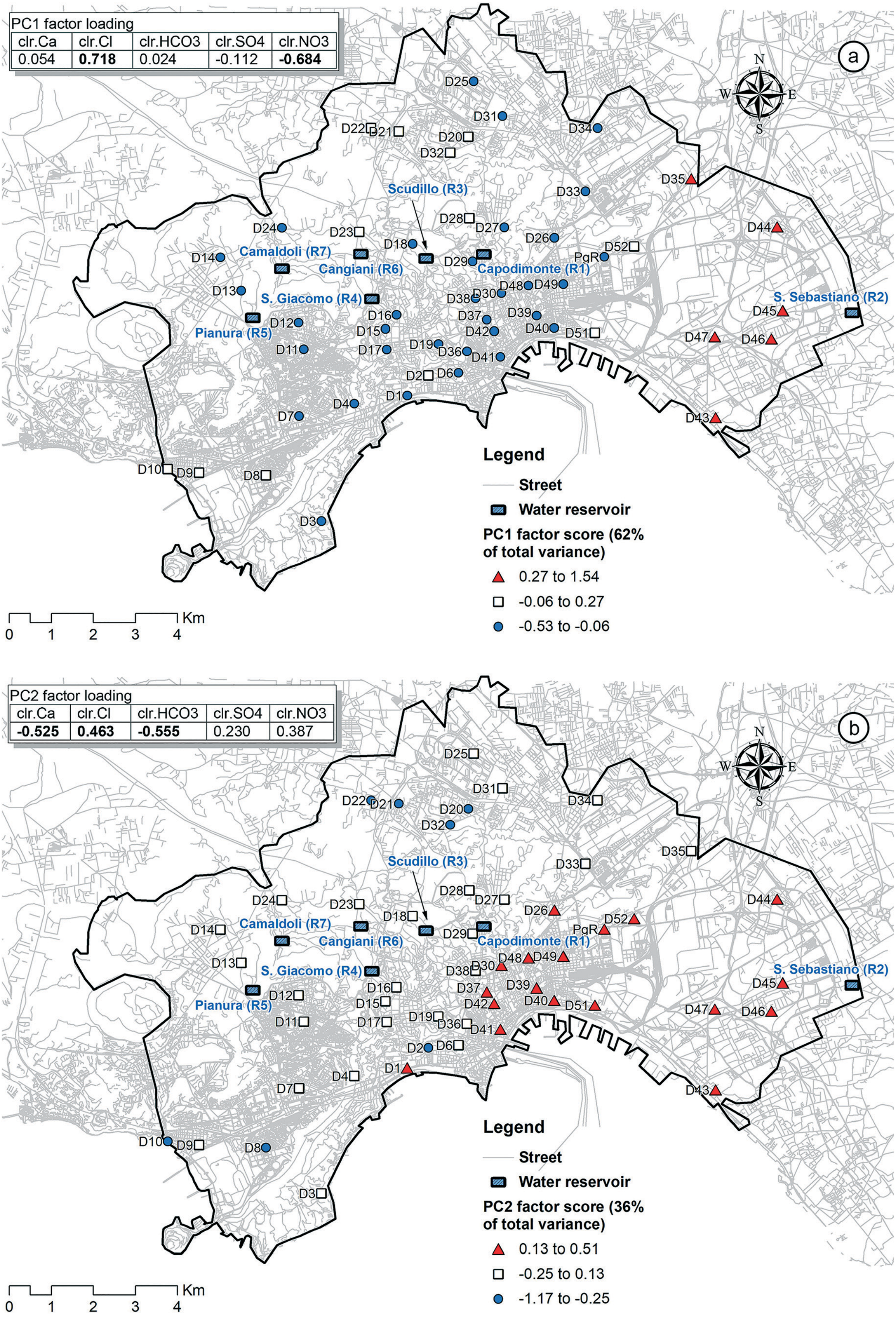
Hierarchical cluster analysis of the tap water samples collected in Naples was performed and the results are presented in Fig. 10. The seasonal differences in the groups suggest that some urban water reservoirs cannot be considered in isolation as single water bodies. The water samples from Capodimonte (R1) and ACO (S1) are clustered into distinct categories indicating the significant role of the water from AIR wells (I5) in the composition of water from the Capodimonte reservoir. In comparison with the other seasons, the samples from Capodimonte and Scudillo (R1 and R3, respectively) are subgroups of the same cluster in spring indicating that Capodimonte makes the greatest contribution to the amount of water in the Scudillo reservoir in this season (Fig. 10b).

Depending on the seasonal interconnections between the urban water reservoirs, the water provided by S. Giacomo (R4), Pianura (R5), Cangiani (R6) and Camaldoli (R7) can be chemically similar to the water from the Serino Aqueduct or Scudillo (Fig. 10). Considering cluster D of HCA in spring (Fig. 10b), the Serino Aqueduct is the main source of the tap water supplied by the above-mentioned reservoirs (S. Giacomo, Pianura, Cangiani and Camaldoli). It seems that the volume of water transferred from Scudillo to S. Giacomo (Fig. 2) gradually increases in winter, summer and fall (Fig. 10). Pianura is the first reservoir that receives Scudillo water in winter (all in cluster B) when there is a low rate of pumping water from Scudillo to S. Giacomo (Fig. 10a). By raising water input from Scudillo in summer, the waters from S. Giacomo, Pianura and Scudillo are grouped in the same cluster owing to their similar chemical characteristics (Fig. 10c). S. Giacomo, Pianura, Cangiani and Camaldoli tap waters (R4, R5, R6 and R7, respectively) are identical with Scudillo tap water (R3) in fall when the water contribution from Scudillo is maximum (Fig. 10d). The above-mentioned discussion might suggest that the Scudillo reservoir is connected to the Pianura reservoir because it is the first one affected by Scudillo water in winter and it occurred in the cluster of Scudillo in three seasons.

Whenever excess water from the Serino Aqueduct (I5) is available, it is mixed with the water in the S. Sebastiano reservoir (R2) to improve the water quality (Fig. 2). In summer, S. Sebastiano probably receives no water from the Serino Aqueduct and consequently most of the R2 and S2 tap water samples are clustered together. However, D35 (together with D43 and D47 in fall) is clustered with group E in winter, spring and fall (Fig. 10). It seems that the pipeline transferring the excess water from the Serino Aqueduct to the S. Sebastiano reservoir (Fig. 2) is not directly connected to the reservoir since transformed values of the ions and physicochemical parameters in D35 (in all seasons except for summer) along with D43 and D47 (in fall) are between the corresponding values of the S. Sebastiano reservoir and Serino Aqueduct (Fig. 11). The non-transformed values of the parameters are also some pieces of supporting evidence (Fig. 7 and S4†).

Except for cluster B in spring where the samples from Capodimonte (R1) and Scudillo (R3) are agglomerated, D6 is never placed in a group with the majority of samples from Capodimonte (Fig. 10). D33 is also clustered with Scudillo instead of R2 and S2 in different seasons. In addition, D27 and D34 do not occur in the group of the Serino Aqueduct (S3). It is consequently proposed that the source of tap water in D6, D27 and D33 is the Scudillo reservoir; however, D34 is representative of S. Giacomo (including Pianura, Cangiani and Camaldoli). On the other hand, D17 which is supposed to receive water from Scudillo is statistically similar to S. Giacomo, Pianura, Cangiani and Camaldoli. It is in accordance with the box plots (Fig. 7 and S4†) and 3D scatterplots (Fig. 11) depicting the chemistry of the statistically analyzed variables and samples.

The chemical fingerprint of the PgR sample which does not exist in the tap-water supply scheme is analogous to the water from the Capodimonte reservoir (Fig. 7, 10, 11 and



PC1 factor loading

clr.Ca	clr.Cl	clr.HCO3	clr.SO4	clr.NO3
0.054	**0.718**	0.024	-0.112	**-0.684**

PC2 factor loading

clr.Ca	clr.Cl	clr.HCO3	clr.SO4	clr.NO3
-0.525	**0.463**	**-0.555**	0.230	0.387

Fig. 9 Factor score maps for the first two principal components explaining 98% of the total variance in the dataset: (a) the first principal component (PC1); and (b) the second principal component (PC2). In order to find an interpretable classification, the upper and lower values of each class were adjusted to match the spatial distribution of different tap waters with the tap-water supply scheme (Fig. 2).

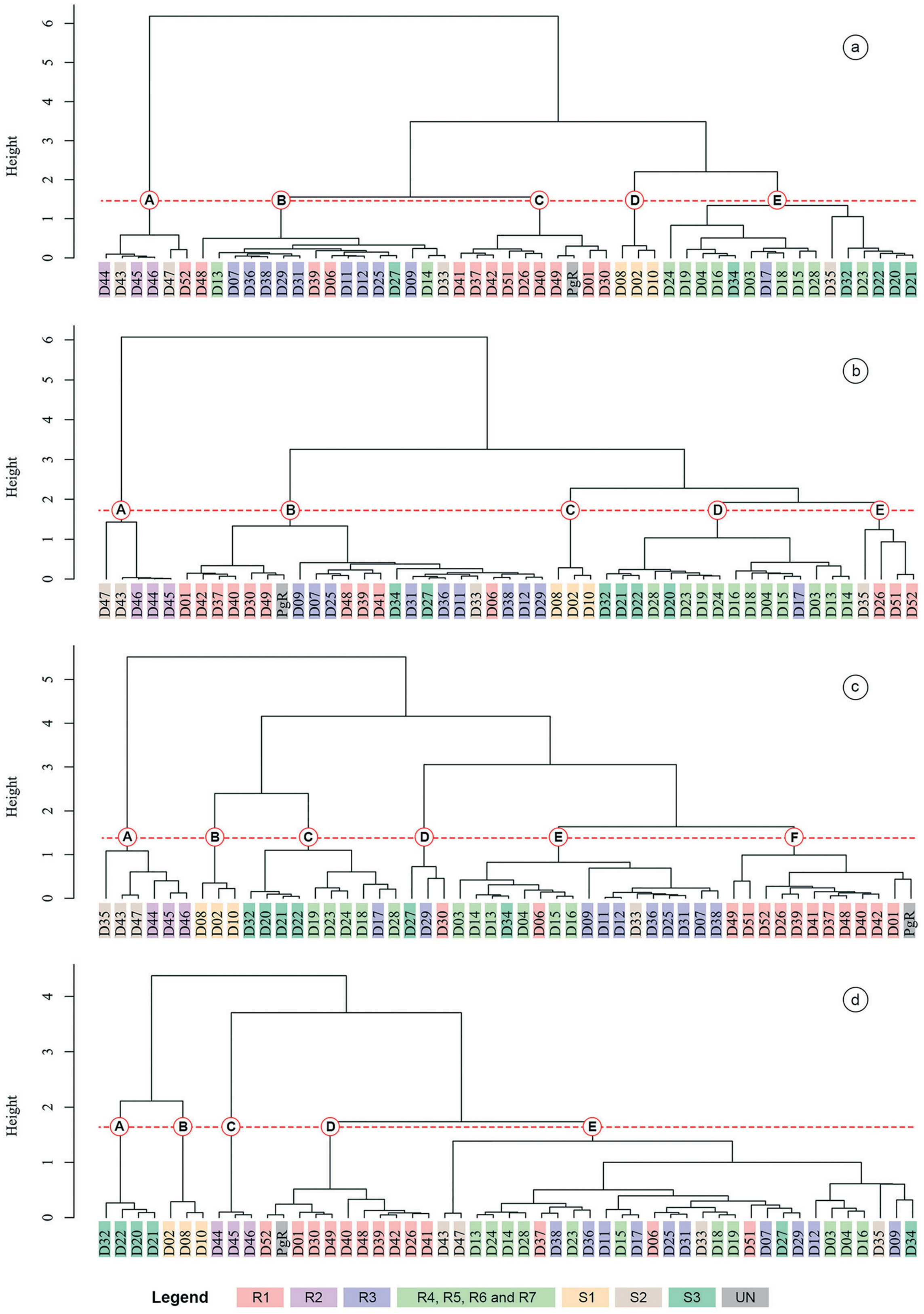


Fig. 10 Hierarchical cluster analysis of the tap water samples using the clr-transformed data: (a) winter; (b) spring; (c) summer; and (d) fall. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme.

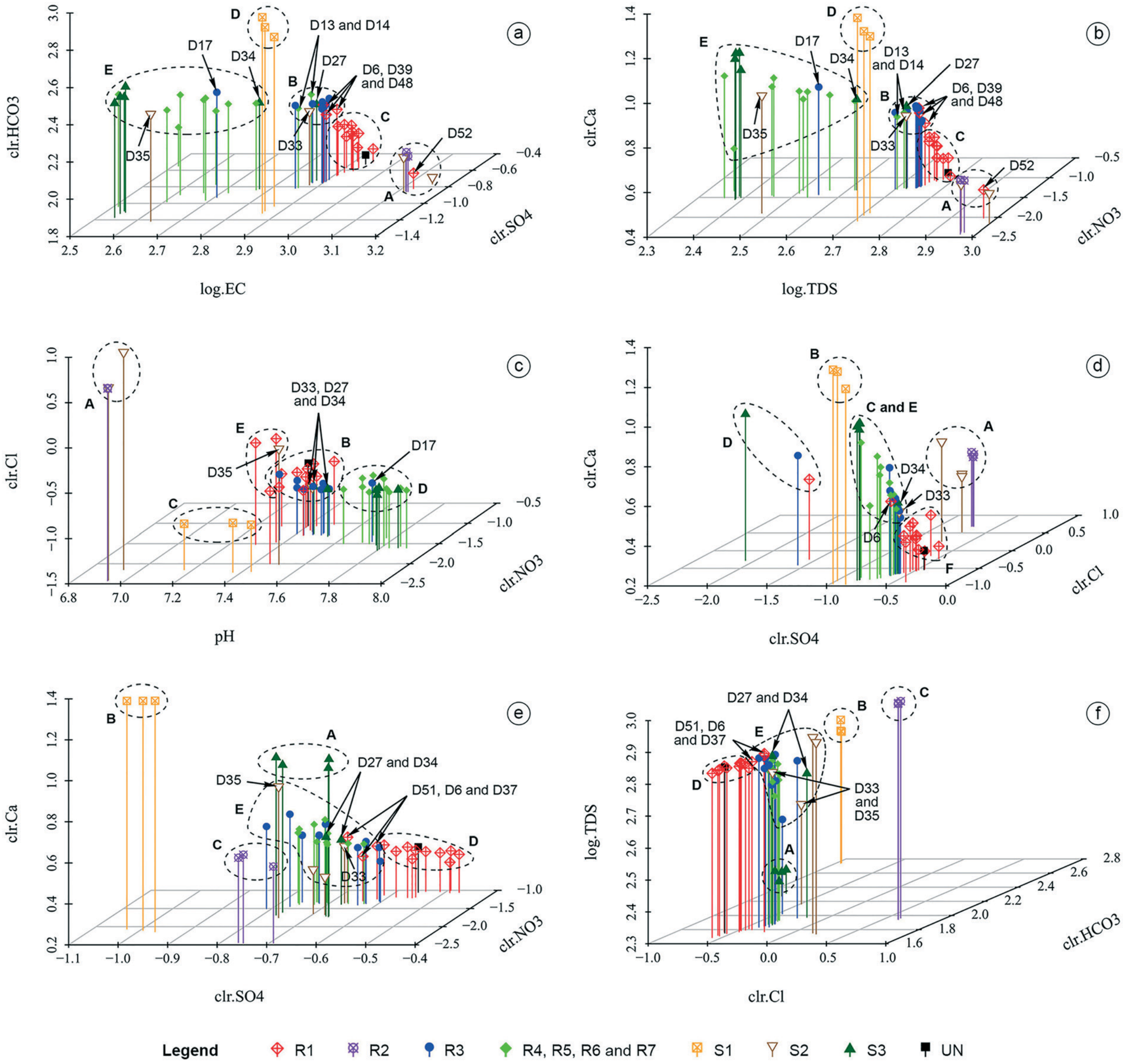


Fig. 11 3D scatter plots of the selected variables: (a and b) winter; (c) spring; (d) summer; and (e–f) fall. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; S3: Serino Aqueduct; and UN: not defined in the water supply scheme. The capital letters (A to E) represent the clusters in Fig. 10.

S4†). The similarity of Ca^{2+}, HCO_3^-, SO_4^{2-}, NO_3^-, EC and TDS transformed values in D35 with those of the Serino Aqueduct, S. Giacomo, Cangiani and Camaldoli led to the occurrence of D35 in group E in winter. Nevertheless, in spring, D35 is clustered with D26, D51 and D52 because of almost the same Cl^- and pH values. The presence of D27, D29 and D30 in group D in summer can be verified by their more or less equal "clr.Cl" and "clr.SO_4" (Fig. 11). Considering the transformed values of Ca^{2+}, HCO_3^-, Cl^-, NO_3^- and TDS in fall, it makes sense that D43 and D47 occur in cluster E with the tap waters from the Serino Aqueduct, S. Giacomo, Pianura, Cangiani and Camaldoli (Fig. 10 and 11).

The information in Fig. 10 is spatially presented by considering some points (Fig. S5†): 1) interpretation of HCA in the previous section is taken into account to avoid the unreliable information in the water supply scheme, 2) with the exception of group E in spring and group D in summer, the name given to each cluster follows the majority of the samples in the group, 3) whenever tap water samples from the Serino Aqueduct occur with S. Giacomo, Pianura, Cangiani and Camaldoli in the same cluster, the cluster is called the Serino Aqueduct which is the main water source, and 4) the group including water samples from Scudillo, S. Giacomo, Pianura, Cangiani and Camaldoli is called Scudillo because of water transfer from Scudillo to the aforementioned reservoirs. It is evident that depending on water transport from one reservoir to another, seasonal maps for water distribution should be considered (Fig. S5†). However, all the four maps are combined to generate a unique map showing the city areas with a distinct source of

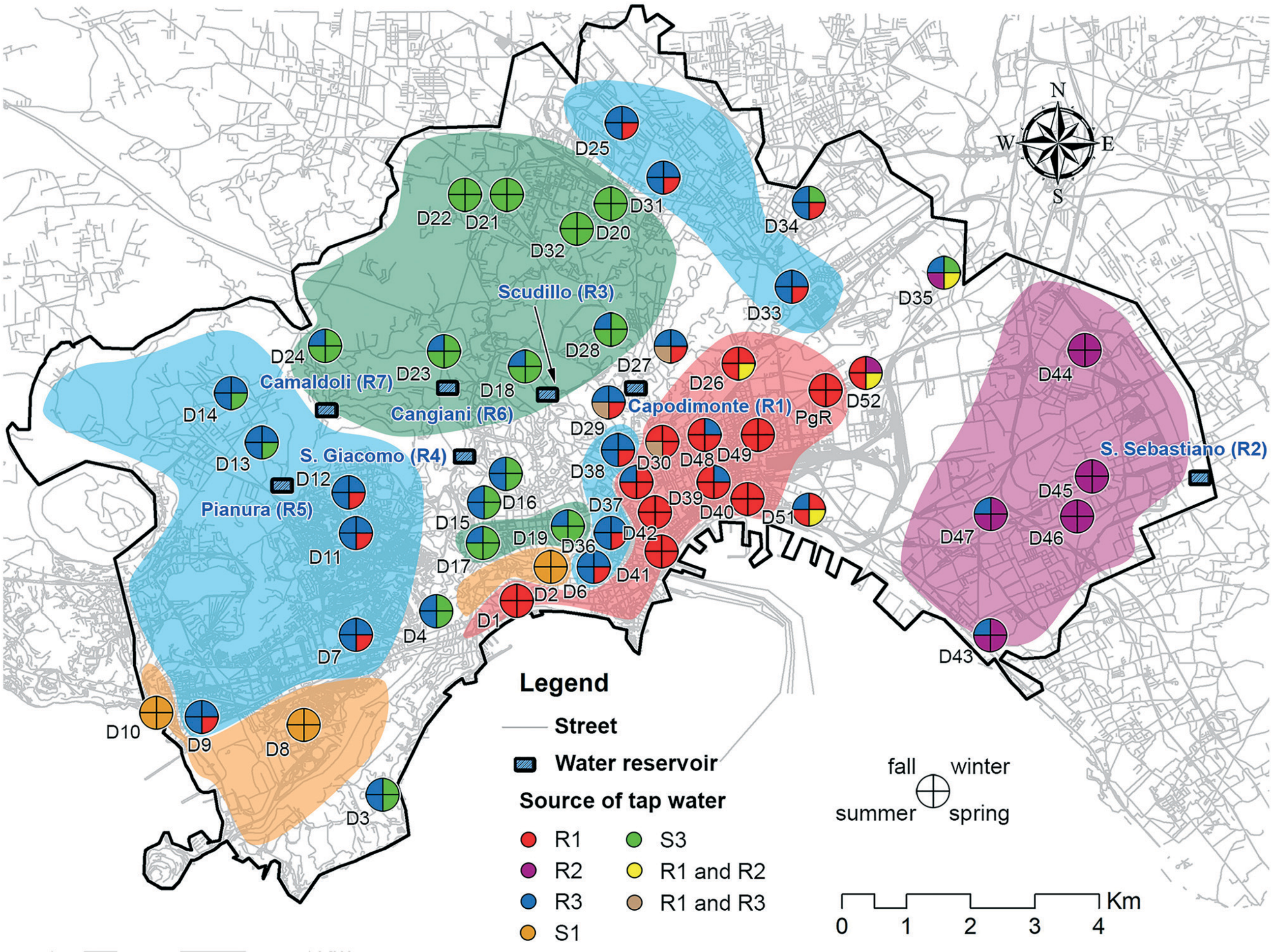


Fig. 12 Results of seasonal HCA indicating the city areas with the same source of tap water. R1: Capodimonte; R2: S. Sebastiano; R3: Scudillo; R4: S. Giacomo; R5: Pianura; R6: Cangiani; R7: Camaldoli; S1: ACO; S2: Cancello well field, Campano Aqueduct and Ponte Tavano wells; and S3: Serino Aqueduct. Color of each city area should be matched with the same color in the legend to define the source of tap water.

tap water (Fig. 12). City areas are identified by the samples with the same HCA result in at least three seasons. In the other parts without zonation, seasonal changes should be considered for better water management. Boxplots of the monthly measurements (Fig. 7 and S4†) and spatial representation of the seasonal HCA (Fig. 12 and S5†) reveal that the water with the least content of cations and anions (*i.e.* the water from the Serino Aqueduct) is distributed in the northern parts of Naples. Nevertheless, residents living in the east consume the water with the highest content of dissolved chemicals. The extent of the S. Sebastiano area (R2 in Fig. 12) is almost equivalent to the city area considered by Giugni *et al.*[20] for hydraulic simulation of the "Napoli Est" water distribution system. In the DMA design, the WDS is usually divided into more homogenous parts with integration of visual examination and technical experience. However, it is complicated in large WDSs with numerous user nodes and pipes.[6] It might explain that the proposed approach can provide some ideas for engineers and managers to divide the complex pipe networks into unique areas prior to hydraulic modelling. The discussion in Text S1† also supports the effectiveness of the methodology presented in this study. Considering the high maintenance operations and the noticeable water loss in the eastern sector of the city,[20] it is highly important to find a solution and manage the WDS in the study area because the problem is feasibly more significant in the historic center of Naples. If it is possible to identify the system's properties (*e.g.*, the network structure and the interactions between its components) in at least some areas defined in the current investigation (Fig. 12), setting up some modelling techniques with minimum prerequisites can further clarify the efficiency of the proposed methodology.

Conclusion

The obtained results show how the chemical data collected for performing periodical water quality checks can be used to assess the flow dynamics in a complex water distribution system fed by multiple supply sources with different geochemical fingerprints. Integrating the principle of CoDA and multivariate statistics for data treatment can be a valuable tool to seasonally monitor the prevalence of the different supply sources in a city like Naples where the urban distribution network has a historical stratified development and a large variety of materials used for constructing pipelines and urban reservoirs. Using the data acquired in different seasons to run the analysis in a wide timespan can represent an opportunity to have a comprehensive view of the distribution network dynamics favoring also the discovery of network interconnections not easily verifiable in the field. In short periods, development of multivariate analysis (especially HCA) using tap water chemistry can lead to a clearer time-specific view of the water distribution in the network together with more effective real-time interventions to locally improve water quality or to locate and manage pipeline failures and systemic water loss. The approach presented in this paper can be used in other complex urban water networks because of the limited cost and positive outcomes. Applying the proposed technique in more case studies and comparing the results with the hydraulic modelling techniques can encourage the scientific community and managers to consider it in future investigations. The performance of the technique can obviously be improved by increasing the sampling density.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors would like to acknowledge Dr. Elena Bottillo, responsible for the Laboratorio Controllo Acque at ABC Napoli, for providing the information on the regional water supplies and the urban distribution network. Further, we are grateful to Dr. Stefano Capasso, the former master student at the Department of Earth, Environmental and Resources Sciences, University of Naples Federico II, who prepared the preliminary version of the dataset. Thanks are also extended to Alberto Fortelli, responsible for the weather monitoring network of RTF Sistemi s.r.l. (www.campanialive.it), for the provision of monthly weather conditions during 2017.

References

1 L. Perelman and A. Ostfeld, Topological clustering for water distribution systems analysis, *Environ. Model. Softw.*, 2011, **26**(7), 969–972.

2 R. Farmani and D. Butler, Implications of Urban Form on Water Distribution Systems Performance, *Water Resour. Manag.*, 2014, **28**(1), 83–97.

3 L. A. Rossman, *EPANET 2 User Manual*, U.S. Environmental Protection Agency (USEPA), Washington, D.C., EPA/600/R-00/057, 2000.

4 A. G. Seyoum and T. T. Tanyimboh, Integration of Hydraulic and Water Quality Modelling in Distribution Networks: EPANET-PMX, *Water Resour. Manag.*, 2017, **31**(14), 4485–4503.

5 J. Izquierdo, M. Herrera, I. Montalvo and R. Pérez-García, Agent-based division of water distribution systems into distric metered areas, in *Proceedings of the 4th International Conference on Software and Data Technologies*, SciTePress - Science and and Technology Publications, 2009, pp. 83–90.

6 M. Scibetta, F. Boano, R. Revelli and L. Ridolfi, Community detection as a tool for district metered areas identification, in *Procedia Engineering*, 2014, pp. 1518–23.

7 H. Bartolín, F. Martínez and J. A. Cortés, Topological GIS-based analysis of a water distribution network model. Applications of the minimum spanning tree, in *Proceedings of the 8th International Conference on Computing and Control for the Water Industry*, CCWI, Exeter, 2005.

8 V. G. Tzatchkov, V. H. Alcocer-Yamanaka and V. Bourguett Ortíz, Graph theory based algorithms for water distribution network sectorization projects, in *8th Annual Water Distribution Systems Analysis Symposium*, American Society of Civil Engineers, Cincinnati, Ohio, USA, 2006.

9 J. W. Deuerlein, Decomposition Model of a General Water Supply Network Graph, *J. Hydraul. Eng.*, 2008, **134**(6), 822–832.

10 J. Deuerlein, S. Elhay and A. R. Simpson, Fast Graph Matrix Partitioning Algorithm for Solving the Water Distribution System Equations, *J. Water Resour. Plan. Manag.*, 2016, **142**(1), 04015037.

11 J. Sempewo, A. Pathirana and K. Vairavamoorthy, Spatial analysis tool for development of leakage control zones from the analogy of distributed computing, in *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference, WDSA*, Kruger National Park, American Society of Civil Engineers, 2008, pp. 676–690.

12 A. Di Nardo and M. Di Natale, A design support methodology for district metering of water supply networks, in *Proceedings of the 12th Annual Conference on Water Distribution Systems Analysis, WDSA*, American Society of Civil Engineers, Tucson, AZ, 2010, pp. 870–887.

13 M. Herrera, S. Canu, A. Karatzoglou, R. Pérez-García and J. Izquierdo, An approach to water supply clusters by semi-supervised learning, in *International Environmental Modelling and Software Society*, iEMSs, Ottawa, 2010.

14 A. Yazdani and P. Jeffrey, Complex network analysis of water distribution systems, *Chaos*, 2011, **21**(1), 016111.

15 M. Scibetta, F. Boano, R. Revelli and L. Ridolfi, Community detection as a tool for complex pipe network clustering, *Europhys. Lett.*, 2013, **103**(4), 48001.

16 K. Diao, Y. Zhou and W. Rauch, Automated Creation of District Metered Area Boundaries in Water Distribution Systems, *J. Water Resour. Plan. Manag.*, 2013, **139**(2), 184–190.

17 O. Giustolisi and L. Ridolfi, New Modularity-Based Approach to Segmentation of Water Distribution Networks, *J. Hydraul. Eng.*, 2014, **140**(10), 04014049.

18 K. Diao, R. Farmani, G. Fu, M. Astaraie-Imani, S. Ward and D. Butler, Clustering analysis of water distribution systems: Identifying critical components and community impacts, *Water Sci. Technol.*, 2014, **70**(11), 1764–1773.

19 Ministro dei Lavori Pubblici, *Decreto Ministeriale 08 gennaio 1997, n. 99: Regolamento sui criteri e sul metodo in base ai quali valutare le perdite degli acquedotti e delle fognature*, Gazz. Uff., 1997, vol. 90.

20 M. Giugni, N. Fontana, D. Portolano and D. Romanelli, A DMA Design for "Napoli Est" Water Distribution System, in *Proceedings of the 13th World Water Congress*, 2008.

21 R. Padulano, Giudice G. Del, M. Giugni, N. Fontana and G. S. D. Uberti, *Identification of Annual Water Demand Patterns in the City of Naples, Proceedings*, 2018, vol. 2(11).

22 Repubblica Italiana, *Decreto Legislativo 2 febbraio 2001, n. 31: Attuazione della direttiva 98/83/CE relativa alla qualità delle acque destinate al consumo umano*, Gazz. Uff., 2001, vol. 52.

23 European Union, Council Directive 98/83/EC of 3 November 1998 on the quality of water intented for human consumption, *Off. J. Eur. Communities: Legis.*, 1998, **3**(L 330), 32–54.

24 M. Civita, Idrogeologia del massiccio del Terminio-Tuoro(Campania), *Memorie e note dell'Istituto di Geologia applicate dell'Università di Napoli*, 1969.

25 P. Celico, L. Esposito, M. De Gennaro and E. Mastrangelo, La falda ad orinte della città di Napoli: idrodinamica e qualità delle acque, *Geol. Rom.*, 1994, **30**, 653–660.

26 D. Ducci and G. Tranfaglia, Effects of climate change on groundwater resources in Campania(southern Italy), *Geol. Soc. Spec. Publ.*, 2008, **288**(1), 25–38.

27 D. K. Todd, *Groundwater hydrology*, Wiley, New York, 1980, p. 552.

28 R. Somma, P. Ebrahimi, C. Troise, G. De Natale, A. Guarino and D. Cicchella, *et al.* 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*, 2021, **274**, 129955.

29 M. Thiombane, J. A. Martín-Fernández, S. Albanese, A. Lima, A. Doherty and B. De Vivo, Exploratory analysis of multi-element geochemical patterns in soil from the Sarno River Basin(Campania region, southern Italy) through compositional data analysis(CODA), *J. Geochem. Explor.*, 2018, **195**, 110–120.

30 J. Palarea-Albaladejo, Martín-Fernández JA. A modified EM alr-algorithm for replacing rounded zeros in compositional data sets, *Comput. Geosci.*, 2008, **34**(8), 902–917.

31 In *Compositional Data Analysis: Theory and Applications*, ed. V. Pawlowsky-Glahn and A. Buccianti, John Wiley & Sons, Chichester, 2011, p. 400.

32 J. Aitchison, The Statistical Analysis of Compositional Data, *J. R. Stat. Soc. Series B Stat. Methodol.*, 1982, **44**(2), 139–177.

33 J. Aitchison, *The Statistical Analysis of Compositional Data*, Chapman & Hall, London, 1986, p. 416.

34 J. M. McKinley, K. Hron, E. C. Grunsky, C. Reimann, P. de Caritat and P. Filzmoser, *et al.* The single component geochemical map: Fact or fiction?, *J. Geochem. Explor.*, 2016, **162**, 16–28.

35 C. D. Lloyd, V. Pawlowsky-Glahn and J. J. Egozcue, Compositional Data Analysis in Population Studies, *Ann. Assoc. Am. Geogr.*, 2012, **102**(6), 1251–1266.

36 X. Gong and M. B. Richman, On the application of cluster analysis to growing season precipitation data in North America east of the Rockies, *J. Clim.*, 1995, **8**(4), 897–931.

37 C. Güler, G. D. Thyne, J. E. McCray and A. K. Turner, Evaluation of graphical and multivariate statistical methods for classification of water chemistry data, *Hydrogeol. J.*, 2002, **10**(4), 455–474.

38 B. Keshavarzi, P. Ebrahimi and F. A. Moore, GIS-based approach for detecting pollution sources and bioavailability of metals in coastal and marine sediments of Chabahar Bay, SE Iran, *Chem. Erde*, 2015, **75**(2), 185–195.

39 J. Aitchison, Principal component analysis of compositional data, *Biometrika*, 1983, **70**(1), 57–65.

40 J. A. Martín-Fernández, in *Compositional Data Analysis*, ed. S. Thió-Henestrosa, Springer International Publishing, Cham, 2016, p. 209.

41 J. Aitchison and M. Greenacre, Biplots of compositional data, *J. R. Stat. Soc. Ser. C Appl. Stat.*, 2002, **51**(4), 375–392.

42 D. Zuzolo, D. Cicchella, A. Lima, I. Guagliardi, P. Cerino and A. Pizzolante, *et al.* Potentially toxic elements in soils of Campania region(Southern Italy): Combining raw and compositional data, *J. Geochem. Explor.*, 2020, **213**, 106524.

43 G. Sawyer, D. McMcarty and G. Parkin, *Chemistry for environmental engineering and science*, McGraw Hill Education, New York, 5th edn, 2003, p. 768.

44 C. Alonso, M. Castellote, I. Llorente and C. Andrade, Ground water leaching resistance of high and ultra high performance concretes in relation to the testing convection regime, *Cem. Concr. Res.*, 2006, **36**(9), 1583–1594.

45 H. Taylor, *Cement Chemistry*, Thomas Telford Publishing, London, 2nd edn, 1997, p. 459.

46 B. Gérard, Hydrolysis of cement-based materials: A review, in *Pore Solution in Hardened Cement Paste*, ed. M. J. Setzer, Aedificatio Publishers, Freiburg, 2000, pp. 271–304.

47 A. Hidalgo, C. Andrade and C. Alonso, An accelerated leaching test to evaluate the long term behaviour of concrete in waste disposal, in *5th CANMET/ACI International Conference on Durability of Concrete*, Barcelona, 2000, pp. 373–381.

48 P. Faucon, F. Adenot, M. Jorda and R. Cabrillac, Behaviour of crystallised phases of Portland cement upon water attack, *Mater. Struct.*, 1997, **30**, 480–485.

49 P. Faucon, F. Adenot, J. F. Jacquinot, J. C. Petit, R. Cabrillac and M. Jorda, Long-term behaviour of cement pastes used for nuclear waste disposal: Review of physico-chemical mechanisms of water degradation, *Cem. Concr. Res.*, 1998, **28**(6), 847–857.

50 B. Lagerblad, Texture and chemistry of historic concrete subjected to prolonged hydration, in *Workshop on water in cement paste and concrete hydration and pore structure*, Nordic Concrete Federation, Skagen, 1999, pp. 148–154.

Publication No. 3

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. https://doi.org/10.1016/j.chemosphere.2021.129955.

Chemosphere 274 (2021) 129955

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

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

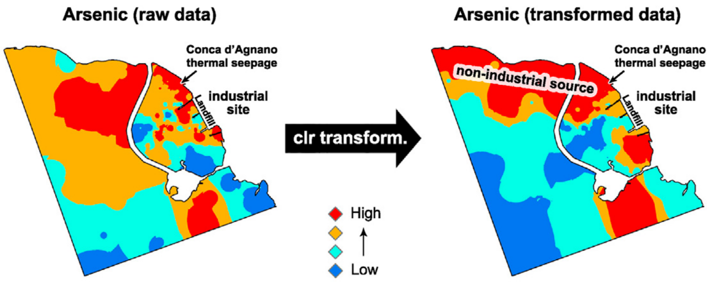
Renato Somma [a, b], Pooria Ebrahimi [c, *], Claudia Troise [a, d], Giuseppe De Natale [a, d], Annalise Guarino [c], Domenico Cicchella [e], Stefano Albanese [c]

[a] *INGV, Osservatorio Vesuviano, 80124, Naples, Italy*
[b] *CNR-IRISS, 80134, Naples, Italy*
[c] *Department of Earth, Environmental and Resources Sciences, University of Naples Federico II, 80126, Naples, Italy*
[d] *CNR-INO, 80078, Pozzuoli, Italy*
[e] *Department of Science and Technology, University of Sannio, 82100, Benevento, Italy*

HIGHLIGHTS

- A transformation was proposed for mapping clr-transformed values with fractal approach.
- Dealing with data closure changes spatial distribution of some PTEs, particularly As and V.
- Hg, Cd, Cu, Pb and Zn were positively correlated with mud and organic matter.
- Association of Al, As, V, Fe, Cr, Ni and sand is stronger along the coast and far from fumaroles.
- Pyroclastic deposits and the seepages with hydrothermal component are the main As sources.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 19 December 2020
Received in revised form
5 February 2021
Accepted 8 February 2021
Available online 11 February 2021

Handling Editor: Willie Peijnenburg

Keywords:
Campi Flegrei
Bagnoli industrial site
Symmetric coordinate
Isometric log-ratio

ABSTRACT

In the last decades, investigating geochemistry of sea sediments has been challenging in the eastern sector of Pozzuoli Bay, source of the metal(loid)s has been a matter of debate and the proposed origin of potentially toxic elements (PTEs) has been occasionally inconsistent. In this study, compositional data analysis (CoDA) was used because the results are independent of the measurement unit, the selected subgroup of elements and the order of chemicals in the dataset. The robust variant of principal component analysis (PCA) indicated that Hg, Cd, Cu, Pb and Zn were positively correlated with mud and organic matter in the sediments deposited in front of the former industrial site. Concentrations of these metals decrease along the cores and in the distal zone. Nevertheless, Al, As, V, Fe, Cr, Ni and sand form an association along the coast which strengthens with increasing distance from fumaroles in the proximal zone. It suggests that arsenic was mainly originated from the pyroclastic deposits of Campi Flegrei and some of the seepages with hydrothermal component, supported by low contribution of the variables in

---

[*] Corresponding author.
E-mail addresses: pooria.ebrahimi@unina.it, pooria.ebrahimi@gmail.com (P. Ebrahimi).

https://doi.org/10.1016/j.chemosphere.2021.129955

Robust principal component analysis
Fractal analysis

robust PCA of the sediments from distal zone. Therefore, this pioneering article suggests CoDA as a powerful tool for answering the long-lasting questions over sediment geochemistry in polluted areas.

1. Introduction

There has been a growing evidence of global environmental damage since the beginning of industrialization and urbanization (Beck et al., 2020). Volcanism and the related activities are also capable of bringing potentially toxic elements (PTEs) to the surface from deep within the Earth (Selinus, 2013). Coexistence of these features in a coastal area can lead to some complexities in geochemical investigations and characterizing source of pollution because sediments can be a source or a sink of chemicals (Keshavarzi et al., 2015; Beck et al., 2020). The restricted water circulation and presence of several industries along the coastline lead to high vulnerability of the Mediterranean Sea to anthropogenic activity. In the late 1980s, the pioneering studies reported industrial pollution of marine sediments in the eastern sector of Pozzuoli Bay (Damiani et al., 1987; Sharp and Nardi, 1987). It raised concerns towards ecological and human health risks of the pollutants and many researchers tried to find source and fate of chemicals by collecting beach sediments, seafloor surface sediments and sediment cores (ICRAM, 2004; Bergamin et al., 2005; Romano et al., 2004; 2009; Albanese et al., 2010; Trifuoggi et al., 2017; Armiento et al., 2020; Daliri et al., 2020; Giglioli et al., 2020). Bergamin et al. (2005) documented positive correlation of Cu, Cr, Mn and Zn with the number of deformed benthic foraminifera. Significant contamination of Cd, Cu, Hg, Mn, Pb and Zn in the seafloor surface sediments collected about 15 years after closure of the local industrial plant indicated that the industrial site is still a source of pollution (Romano et al., 2009).

In the last decade, Albanese et al. (2010) illustrated that the metal-rich groundwater which mainly originates from hydrothermal activity of Campi Flegrei volcano has carried the pollutants from the coastal landfill to the marine environment. The highest As concentrations showed a strong spatial correlation with the submarine thermal springs. Arsenic contents decreased with increasing seawater depth (Somma et al., 2016) because hot hydrothermal fluids probably limit As precipitation (Albanese et al., 2010). Trifuoggi et al. (2017) dredged seafloor surface sediments of Pozzuoli Bay along five transects perpendicular to the coastline and showed that the average concentrations of As, Cd, Hg, Pb and Zn were two to twenty times greater than the Italian regulatory guidelines. They concluded that the insignificant correlation between the mean grain size and PTEs (i.e., As, Cd, Cr, Cu, Hg, Ni, Pb and Zn) excludes terrigenous sediments and weathering products as their main sources.

Recently, Armiento et al. (2020) showed that As, V, Cr and Ni were controlled by geogenic (volcanism and the related fumaroles for As and V, but minerals for Ni and Cr) and anthropogenic factors. Nevertheless, Giglioli et al. (2020) demonstrated that pyroclastic rocks, groundwater, seepages, fumaroles and industrial activity have introduced As to the coastal environment. While wave hydrodynamics affects spatial distributions of the metals originated from the coastal landfill (i.e. Cd, Cu, Hg, Pb, Zn and Fe) together with As, it does not influence geochemical distributions of Cr, Ni and V mostly concentrated in the sediments deposited offshore and NW Nisida island. Because sediment toxicity is not mostly representative of sediment chemistry, Morroni et al. (2020) used an integrated approach and identified a “Slight” to “Moderate” level of risk, with the highest near the plant.

Regarding the above-mentioned studies, the underwater fumaroles, submarine groundwater discharge, minerals, extensive industrial activity in the past, ineffective disposal of the industrial wastes in the coastal landfill and the seepages were reported as the factors governing sediment geochemistry in the study area. It indicates that there has been a long-lasting debate between researchers about source of some PTEs (particularly As) in the marine sediments due to the complex geochemical environment. Compositional data analysis (CoDA) is getting popular because of its potential to avoid doubtful results. Separation of the geochemical signals of grain size, anthropogenic contamination and possible post-depositional alteration in fluvial sediments of NW Czech was facilitated by utilizing CoDA tools (Álvarez-Vázquez et al., 2020). In another study, Dominech et al. (2020a) investigated cadmium in the sediments collected along Sabato river (south Italy) and reveled capability of CoDA in showing the relative geochemical information at catchment basin scale. The extent of mineralization, weathering, diagenesis, contamination and a combination of these factors can be effectively detected following log-ratio transformation of the analytical data obtained from stream sediments (Wang et al., 2014; Liu et al., 2016, 2018). The idea behind CoDA was also applied for paleoenvironmental reconstruction of sedimentary environments (Flood et al., 2016; Razum et al., 2020).

To the best of our knowledge, log-ratio transformations were only utilized by Wang et al. (2015) to generate single-element geochemical maps in marine sediments. Therefore, based on the CoDA framework, this is the first work that presents the results of multivariate statistical analysis of PTEs in seafloor surface sediments and core sediments. It is expected that considering compositional nature of data leads to some useful information about the origin of elements and reveals the capabilities of this approach for investigating geochemistry of polluted sea sediments. Furthermore, a data transformation will be proposed to enable fractal analysis for geospatial analysis of negative values obtained after log-ratio transformation.

2. Study area

The Bagnoli urban district (Fig. 1) in the city of Naples is located on the eastern coast of Pozzuoli Bay (Tyrrhenian Sea, Italy). For about 100 years, production of steel, asbestos, cement and fertilizer in an industrial site established near the urban area strongly polluted the environment. It is worth mentioning that the steel plant was one of the most important Italian steelworks in the 20th century which reached a maximum of 2×10^6 million tons of steel production in 1960. The modifications of human to the natural environment have been one of the key factors in environmental crisis in the study area. In 1930, two piers were constructed for ship berthing (Fig. S1). The raw materials such as coal and iron ore were unloaded in the northern one although the products were loaded in the southern one. Marine hydrodynamics was accordingly changed, and sediment transport capacity has significantly decreased in front of the northern pier. Five years later, Nisida island was connected to the mainland by a stone bridge which also altered the water circulation pattern and sea sediment movement. In order to construct buildings, store coal and develop the industrial activity,

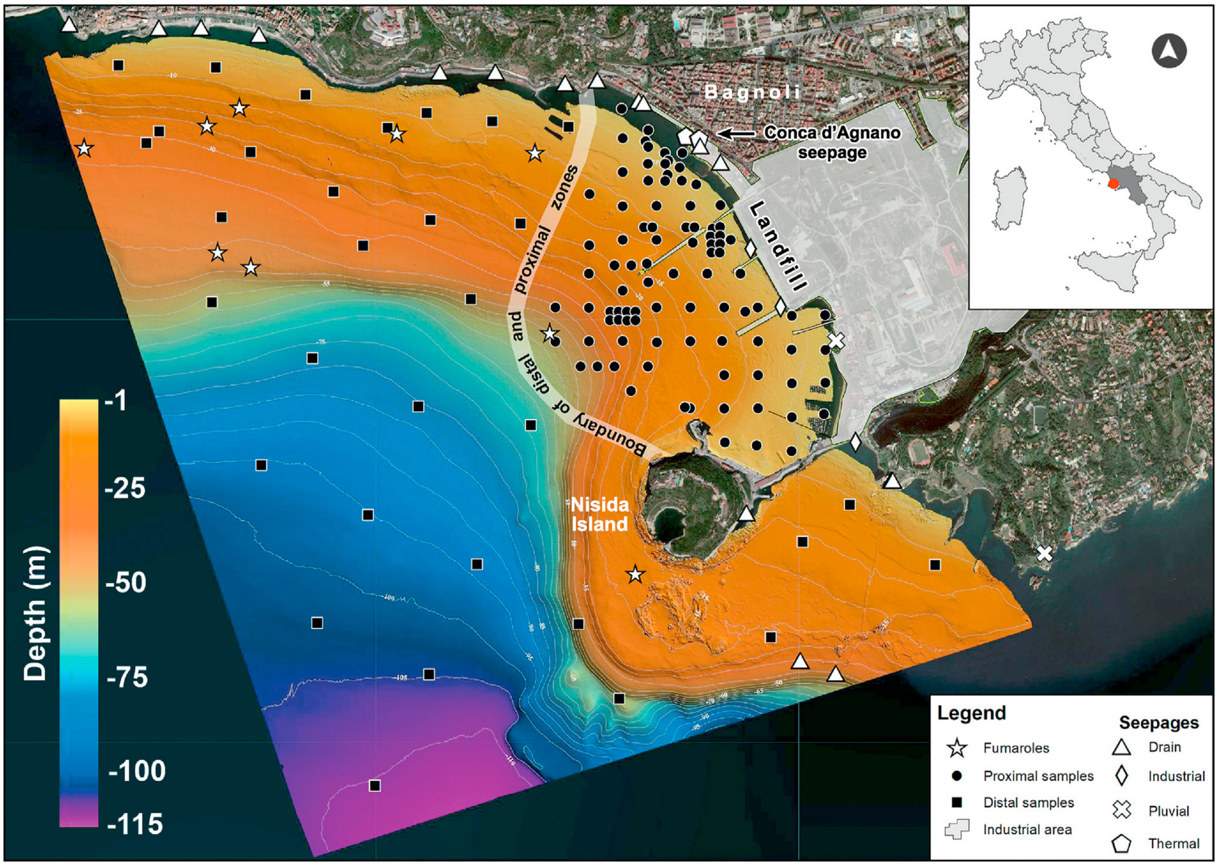


Fig. 1. Location of the study area, bathymetry of the coastal environment, the sediment samples of proximal and distal zones (PZN and DZN, respectively) and location of seepages and fumaroles.

the area between the piers was partly filled with contaminated soil of the industrial site in the early 1960s (Fig. 1) (Romano et al., 2009; Arienzo et al., 2019; Armiento et al., 2020). The shoreline dynamics generally depends on the influence of wind and tide on the emerged beach together with sediment mobilization by wave motion and marine currents in the submerged beach. The winds and waves that approach from southwest and west lead to a NW-SE drift by longshore currents and control the littoral morphodynamics. The effective fetch and depth of closure calculated for 1951–2009 period are 280 km and 7.5 m, respectively (Arienzo et al., 2019).

On the other hand, the study area is a part of the Campania Volcanic Zone (Rolandi et al., 2020) and lies within the active volcanic area of Campi Flegrei caldera (Fig. S1) formed 15 ka BP by a large ignimbrite eruption named Neapolitan Yellow Tuff (De Natale et al., 2016; Troise et al., 2019). Neapolitan Yellow Tuff is covered with a sequence of interdigitated pyroclastic deposits produced by recent volcanic and tectonic activities and then lacustrine and palustrine sediments. The last eruption occurred in 1538 AD, but the intense hydrothermal circulations have resulted in surface manifestations of hot springs, steam-heated pools, together with submarine (Fig. 1) and subaerial fumaroles. The ascending hydrothermal fluids bring considerable loads of PTEs such as arsenic to the surface (Aiuppa et al., 2006).

The above-mentioned discussion reveals that concentrations of PTEs are regulated with anthropogenic and geogenic features. Although closure of the industrial site occurred in 1990s, the area which pose a high environmental risk was included in the list of polluted sites of national interest (SIN) for monitoring and recovering the area. Remediation programs were started in 1996 and extended to the marine sediments in front of the brownfield five years later (Albanese et al., 2010). ABBaCo was the latest project which aimed to (ABBaCo Project, 2018): (1) upgrade the information about environmental quality of Bagnoli industrial site; (2) characterize the pollution sources; and (3) propose appropriate remediation strategies.

3. Materials and methods

3.1. Sampling and chemical analysis

Sampling campaigns were planned in May and November 2017 to collect sediment samples from the proximal (near the Bagnoli industrial area) and distal (far from the industrial area) zones (hereafter PZN and DZN, respectively; Fig. 1) regarding the preliminary investigations of the Italian Higher Institute for Environmental Protection and Research (ISPRA). In the PZN, 93 sediment cores were drilled using a vibro-corer with 10 cm inner diameter. However, a Van Veen grab was used to collect 32 samples from the upper 20 cm in the DZN (Fig. 1). About 1 kg of sediment was placed into a plastic bag and stored at approximately 4 °C until transfer to the laboratory. Coordinates of the sampling stations were obtained via a differential global positioning system.

Laboratory sampling and sediment pre-treatment were performed following the Ministerial Decree 7/11/2008. Each sediment core was accordingly sub-sectioned into 0–50, 50–100, 100–150, 150–200 and 200–300 cm intervals depending on the core depth. The aliquots for metals/metalloids (except for Hg) were oven-dried at 105 °C for 24 h whilst the ones for Hg were air-dried. All samples were subsequently homogenized and stored at 4 °C before laboratory analyses according to the guidelines and directives for assessment of marine pollution in SINs (Legislative Decree, n. 152/06). For grain size analysis, the sediments were successively treated with H_2O_2 solution, washed with distilled water and dried at 40 °C prior to weighing. Samples were then wet-sieved using a 63-μm sieve, and the coarser fraction was further separated into different grain sizes by passing them through ASTM series sieves with meshes ranging between −2 and +4 phi (Romano and Gabellini, 2001). Finally, the muddy (<63 μm), sandy (63 μm–2 mm) and gravelly (>2 mm) fractions were calculated for the purpose of this study.

Total concentrations of elements were measured in the <2 mm grain size by microwave-assisted acid digestion procedure following the EPA 3052 method. The analyses of the metal(loid)s were performed through the following protocols and instruments: (1) EPA 6020b for As, Cd, Cr, Cu, Ni, Pb, V and Zn determination with inductively coupled plasma-mass spectrometry (PerkinElmer Elan-6100); (2) EPA-6010d for Al and Fe measurement with inductively coupled plasma-optical emission spectrometry (PerkinElmer Optima-2000DV); and (3) EPA 7473 for Hg detection with the Automatic solid/liquid Mercury Analyser (FKV AMA-254). The computed quantitation limits for dry weight were as follows (mg/kg): 0.05 for As, V and Pb, 0.03 for Cd, 0.5 for Cr, 0.3 for Cu and Ni, 1.0 for Zn, 50 for Al, 5 for Fe and 0.005 for Hg. In order to check accuracy of the analytical determinations, BCSS-1, MESS-4 and PACS-3 marine sediment standard reference materials were analyzed using the procedures applied to the sediment samples and the obtained recovery rates ranged from 80 to 118% (Table S1).

3.2. Compositional data analysis (CoDA)

In geochemical and environmental datasets, variation of most parameters should be assessed relative to the other variables because only ratios between parts are informative. This is well explained by the fundamentals of CoDA in which a log-ratio methodology transforms the compositional data from their original sample space to an unrestricted real space (Filzmoser et al., 2018). The proposed family of log-ratio transformations includes the additive log-ratio (alr), centered log-ratio (clr) and isometric log-ratio (ilr) transformations (Aitchison, 1982; Egozcue et al., 2003). The values obtained from the first two log-ratios are subject to some restrictions in standard statistical techniques whilst the ilr-transformed data can be investigated by standard univariate and multivariate statistical approaches (Lloyd et al., 2012). In the present study, the elements and the grain size fractions were considered as two different sets of compositional data.

3.2.1. Spatial distribution of single variables: raw data vs. clr-transformed data

In order to obtain a general insight into the coastal sediment geochemistry, the interpolated maps of the raw data and the clr coordinates were generated (Fig. S2). This transformation has been used to decipher the geochemical patterns in sediments by several researchers (e.g., Liu et al., 2016; Dominech et al., 2020a, 2020b). The clr transformation for a composition $x = [x_1,\ x_2, \ldots,\ x_D] \in S^D$ is the transformed data $y \in R^D$ with (Lloyd et al., 2012):

$$y = \mathrm{clr}(x) = \left(\ln \frac{x_1}{\sqrt[D]{\prod_{i=1}^{D} x_i}},\ \ln \frac{x_2}{\sqrt[D]{\prod_{i=1}^{D} x_i}}, \ldots, \ln \frac{x_D}{\sqrt[D]{\prod_{i=1}^{D} x_i}} \right) \quad (1)$$

Computing the clr coefficients is not complicated and assigning each original compositional part to a clr coordinate relatively facilitates interpretation of the transformed values. However, this transformation is sub-compositionally incoherent which may result in misinterpretations when including a subset of variables in the denominator (Razum et al., 2020). Another restriction of the clr coefficients is that they sum to zero and their covariance and correlation matrices are singular (Buccianti et al., 2006). Thus, the ilr transformation (more information is provided in the following sections) was also considered for validation of the geochemical maps produced following clr-transformation and for identification of the main processes regulating geochemistry of PTEs.

Because a non-negative dataset is required for multifractal inverse distance weighted (MIDW) mapping (see Text S1 for details), a data treatment was proposed in this article to deal with the negative clr coefficients (Fig. S2). First, the variables (i.e. Al, Fe and Zn) with positive clr values (*clr.element*) were scaled (clr_{scl1}*.element* and clr_{scl2}*.element*) using equations 2 and 3 in Table 1. Then, the spatial distributions of *clr.element*, clr_{scl1}*.element* and clr_{scl2}*.element* were prepared and reclassified according to the fractal approach. The upper and lower values of each class were subsequently back-transformed to the original *clr.element* by the equations 4 and 5 (Table 1). It was illustrated that applying the proposed formulas to scale the clr coordinates does not significantly affect the spatial patterns (Figs. S3–S5). Therefore, whenever a negative *clr.element* is present in a variable, Eq. (2) is utilized to obtain the threshold concentrations by fractal analysis which will be back-transformed to the original clr coefficients via equation 4. The above-mentioned approach was also utilized for mapping grain size fractions, principal component scores and standardized residuals which will be further discussed in the following sections. The MIDW mapping was performed using the ArcFractal Plugin in ArcGIS 10.4.1 software.

3.2.2. Correlation analysis: the log-transformed data vs. the symmetric coordinates

At the end of nineteenth century, Pearson (1897) explained the problems with correlation analysis applied to the parts of a composition and demonstrated that obtaining reliable results depends on capturing the relative information of compositional parts. The clr transformation is not suitable for this purpose due to the drawbacks discussed in the previous section. Pivot coordinates (a special choice of the ilr transformation) are an alternative, but the direct relationship between the original variables and the transformed values will be lost (Reimann et al., 2017). To overcome the shortcomings of clr and ilr transformations and evaluate the relationship between two compositional parts, construction of the symmetric coordinates as a special case of the ilr transformation is advantageous. It preserves the dominance of both parts in a

Table 1
Scaling the computed clr values (*clr.element*) to obtain positive values for MIDW interpolation. The cut-off scores were subsequently back-transformed to the original *clr.element*.

	Scaling the *clr.element*		Back-transforming to the *clr.element*	
All variables	$clr_{scl1}.element = clr.element - Min_{clr.element} + 0.1$	Eq. (2)	$clr.element = clr_{scl1}.element + Min_{clr.element} - 0.1$	Eq. (4)
Variables with *clr.element* > 0	$clr_{scl2}.element = clr.element + Min_{clr.element} + 0.1$	Eq. (3)	$clr.element = clr_{scl2}.element - Min_{clr.element} - 0.1$	Eq. (5)

composition by generating two pivot coordinates in a symmetric manner (Kynčlová et al., 2017). Further information about calculation of symmetric coordinates is provided in Text S1.

Because the classical log-transformation of single variables prior to evaluating the relationship between parameters is still common, the log-transformed data and the symmetric coordinates were utilized to conduct correlation analysis and demonstrate the importance of dealing with data closure (Fig. S2). The R package “robCompositions” (Filzmoser et al., 2018) was used for transforming the data and generating the heatmaps of Spearman correlation.

3.2.3. Principal component analysis (PCA): the classical vs. robust approaches

The covariance structure of compositional data is significantly biased without an appropriate data transformation and in presence of data outliers (Maronna et al., 2006) which leads to doubtful results of multivariate statistical analysis. Robust estimation of the covariance matrix was proposed as a solution in the robust PCA (Filzmoser, 1999). Contrary to classical PCA (Aitchison and Greenacre, 2002), it does not work with clr-transformed data because a full rank data matrix is usually needed for robust covariance estimators. Therefore, Filzmoser et al. (2009) proposed applying the ilr transformation prior to robust PCA and subsequently back-transforming the resulting scores and loadings to the clr space for better interpretation (Fig. S2). Following PCA, biplots can be used for representing results and showing relationships (see Text S1 for details).

3.2.4. Sediment pollution

Although CoDA is a powerful tool for geochemical investigation (Thiombane et al., 2018; Dominech et al., 2020a, 2020b), the currently available pollution indices do not address the compositional nature of data. To bridge this gap, balances which are a family of the ilr coordinates were constructed through sequential binary partitioning (Egozcue and Pawlowsky-Glahn, 2005). It helps to easily interpret the relative information about two groups of parts that cause the major effects within a composition (see Text S1 for details). Linear regression model was subsequently applied to the balances for modelling the relationship between metal(loid)s and granulometry as dependent and independent variables, respectively. The residuals were then divided by their standard error to obtain the standardized residuals which can be indicators of anthropogenic contamination or geogenic enrichment (Fig. S2) (Petrik et al., 2018). These values were finally mapped via MIDW approach as described in the section 3.2.1. The linear regression model was performed through the R package “robustbase” (Maechler et al., 2020).

4. Results and discussion

4.1. Geochemistry of seafloor surface sediments

This section refers to the sediments of 0–50 cm in PZN and 0–20 cm in DZN.

4.1.1. Physicochemical characteristics of the sediments

The descriptive statistics of grain size fractions and organic matter (OM) contents of the samples are shown in Table S2. In the PZN, the proportion of gravel, sand and mud range from 0–22%, 26–99% and 1–70%, respectively. On the other hand, variation of sediment grain size is higher in the DZN: 0–55% for gravel, 9–99% for sand and 0–91% for mud. On average, sand is the dominant fraction. However, higher variation of mud (CVR>70%) and gravel (CVR>85%) in both datasets might explain the role that wave plays in transporting sediments (Table S2). The sediment type (Schlee, 1973) is mainly sand, silt and clay, but some samples are gravelly sediment (Fig. S6). It is worth mentioning that because elements were determined in the <2 mm size fraction, the gravel was only used for sediment type classification. According to Fig. 2, sand is mainly deposited along the coast whilst mud is more frequent (sometimes the major fraction) in the low-energy environment in front of Bagnoli landfill (De Pippo et al., 2002). The clr-transformation of the data does not change the spatial pattern of coarse and fine sediment fractions (Fig. 2a–d).

The median OM content of the sediments from PZN (2.18%) is approximately twice that of DZN (0.92%). The highest OM values occur near the Bagnoli brownfield site and the levels decrease towards the open sea and the shoreline (Fig. 2e). The fine-grained sediments and OM can be important metal carriers (Ebrahimi, 2013; Sciarrillo et al., 2020), which will be further investigated in the following sections.

4.1.2. Metal(loid) contents of the sediments

Statistical summary of the analyzed metal(loid)s are presented in Table S2. Considering the 95th percentile, concentrations of the elements are greater in the PZN except for Ni, Cr and V with almost equal levels in both zones. The average contents of the elements follow the decreasing order of Fe > Al > Zn > Pb > V > As > Cu > Cr > Ni > Cd > Hg in the PZN and Al > Fe > Zn > Pb > V > As > Cr > Cu > Ni > Cd > Hg in the DZN. Contrary to Al, Fe, As, Ni, Cr and V, the CVR values of Hg, Cd, Pb, Zn and Cu are above 60% in the PZN which likely suggest presence of unusual processes (Reimann et al., 2008). The CVR values of Cd, Pb and Zn decrease in the DZN while those of Hg and Cu change insignificantly.

Evidently, Fig. 3 indicates that there is a considerable difference between spatial distribution of raw data and the clr coordinates of As and V. Similar changes were also observed for Al, Fe, Cr and Ni (Figs. S7 and S8). Thus, it might stress the importance of opening the geochemical data before geospatial analysis. Regarding the MIDW maps (Figs. 3, S7 and S8), the higher clr.As, clr.V, clr.Al, clr.Fe, clr.Cr and clr.Ni values mainly occur along the coastline and do not coincide with the industrial activities. However, spatial patterns of the other chemicals remained almost invariable after log-ratio transformation. In PZN, the sediments characterized by almost high contents of Hg, Cd, Pb, Zn and Cu are deposited in front of the local industrial site and the highest concentrations occur in greater water depth (Fig. 4 and S9). This geochemical pattern can be explained by wave turbulence in front of the brownfield which encourages suspension and transportation of sediments from the emerged and submerged beaches (i.e. from 4.3 m a.s.l. to 6.5 m b.s.l.) due to run-up and depth of closure (Arienzo et al., 2019). Extension of the contaminated plume in the DZN can also be verified through the elevated clr values of these PTEs .

4.1.3. The statistical significance of data transformation

The Spearman correlation between the elements is presented in the proximal and distal sediments (Fig. 5). It is evident that, in agreement with the discussion in the previous section, CoDA clustered the variables into two distinct groups (Fig. 5a and c): (1) Al, As, Cr, Ni, V and Fe; and (2) Hg, Pb, Zn and Cu. Cadmium is correlated with the second group in the sediments near the industrial site, but it is associated with the first group in the sediments far from Bagnoli. In addition, Cr and Ni revealed weaker correlation coefficients with Al, As, V and Fe in the DZN. However, majority of the elements are positively correlated together in the right panels constructed via log-transformed data (Fig. 5b and d) which prevents effective data interpretation. The compositional

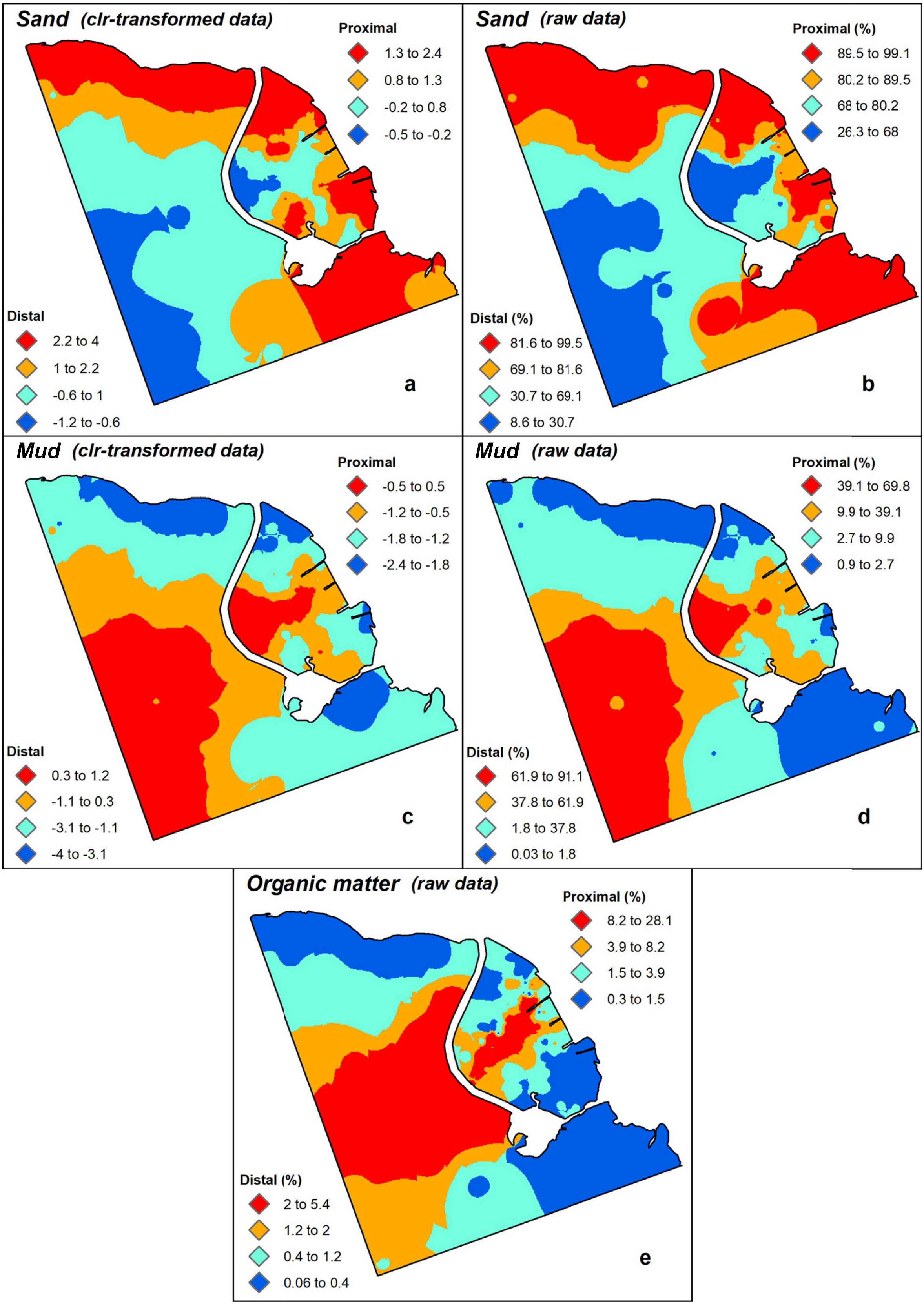


Fig. 2. Spatial distribution of sand, mud and organic matter produced by clr-transformed data (a and c) and raw data (b, d and e). Dealing with data closure does not significantly influence spatial patterns of the grain size fractions.

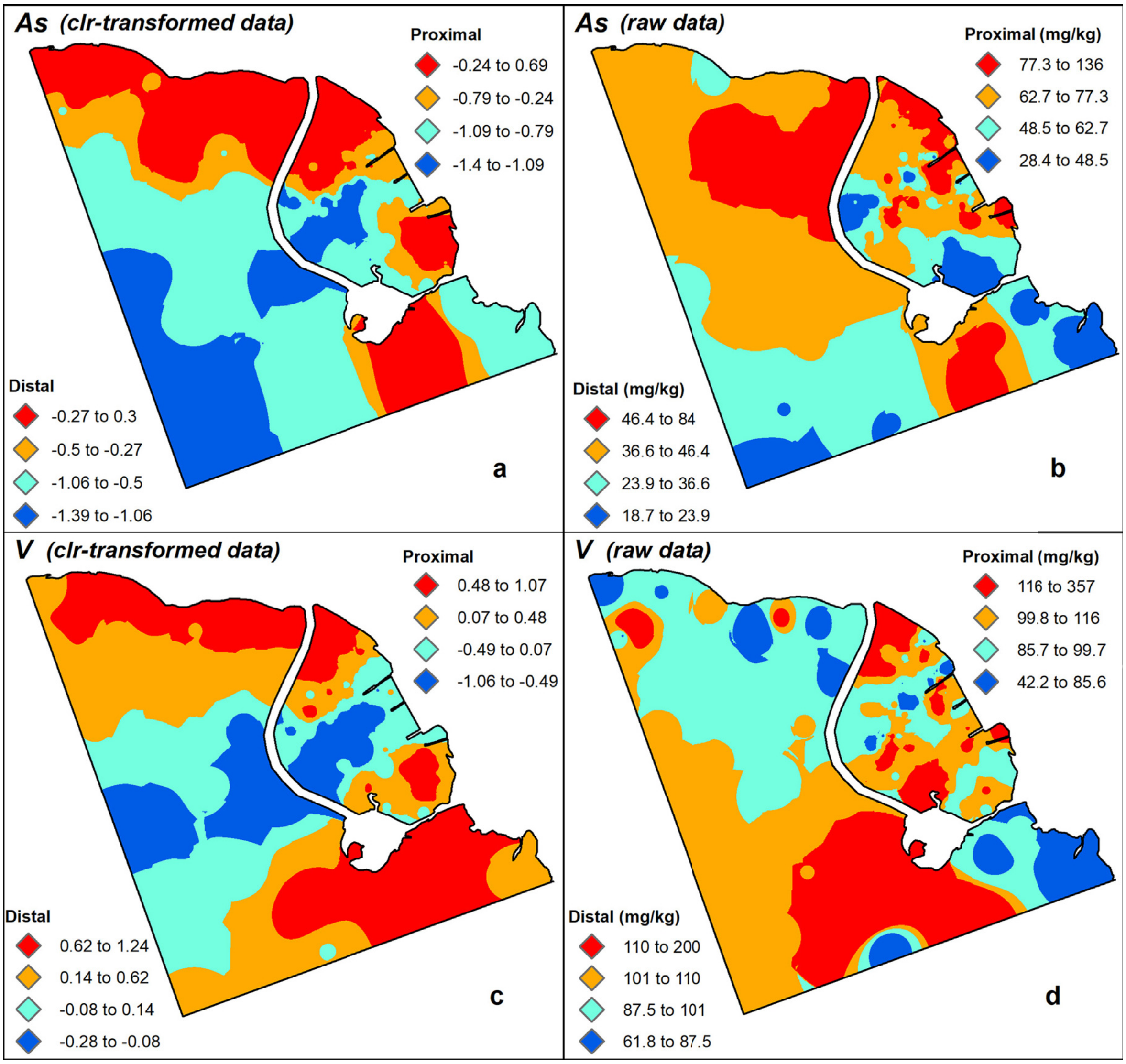


Fig. 3. Arsenic and vanadium geochemical maps produced by clr-transformed data (a and c) and raw data (b and d). Opening the data noticeably impacted spatial patterns of the PTEs.

approach was consequently used in the following sections for reliable data mining and informative pattern recognition.

4.1.4. *The processes governing geochemistry of the sea sediments*

Regarding the results of correlation analysis (Fig. 5), CoDA was integrated with PCA to find the main geochemical processes. The rise of hot thermal fluids through underwater fumaroles (Albanese et al., 2010), minerals, industrial activity (Armiento et al., 2020) together with discharge of As-rich groundwater (Aiuppa et al., 2006; De Vivo et al., 2008) and a mixture of thermal fluids and municipal wastewaters (Giglioli et al., 2020) in the sea were introduced as the probable sources of arsenic in the study area. In addition, the fine fraction of sediments and organic materials are effective metal scavengers (Keshavarzi et al., 2015; Sciarrillo et al., 2020). Therefore, sand and mud were treated as the second set of compositional data whilst organic matter (OM) along with the shortest distance of each sample to fumaroles, coastline and seepages (i.e. dist_fumarole, dist_coast and dist_seepage, respectively) were considered as external non-compositional variables in PCA. Presence of data outliers was verified prior to statistical analysis via scatterplots of the symmetric coordinates of elements (Fig. S10). Hence, the classical and robust PCA were carried out to evaluate impact of the outliers on the outcomes and it was apparent that the robust method had a better performance because it explained a greater variance in the dataset (Figs. 6 and S11). Regarding the proximal sediments, biplot of the first two principal components which accounted for 89.35% of the total variance was very informative and grouped the variables in two associations (Fig. 6a): (1) Hg, Cd, Cu, Pb, Zn, OM, mud, dist_coast and dist_seepage; and (2) Al, As, V, Fe, Cr, Ni, sand and dist_fumarole. On the other hand, 94.73% of variance was explained by PC1 and PC2 in the distal sediments (Fig. 6b). Compared with the PZN, shorter rays of the elements highlight their lower variation in the sediments deposited far from the Bagnoli industrial site. According to the PCA biplot (Fig. 6b), two relevant elemental associations are as follows: (1) Hg, Cu, Zn, Pb, OM, mud, dist_seepage, dist_fumarole and dist_coast; and (2) Al, As, V, Fe, Ni, Cr and sand. Cadmium was excluded because of the lowest variability in PCA. This might

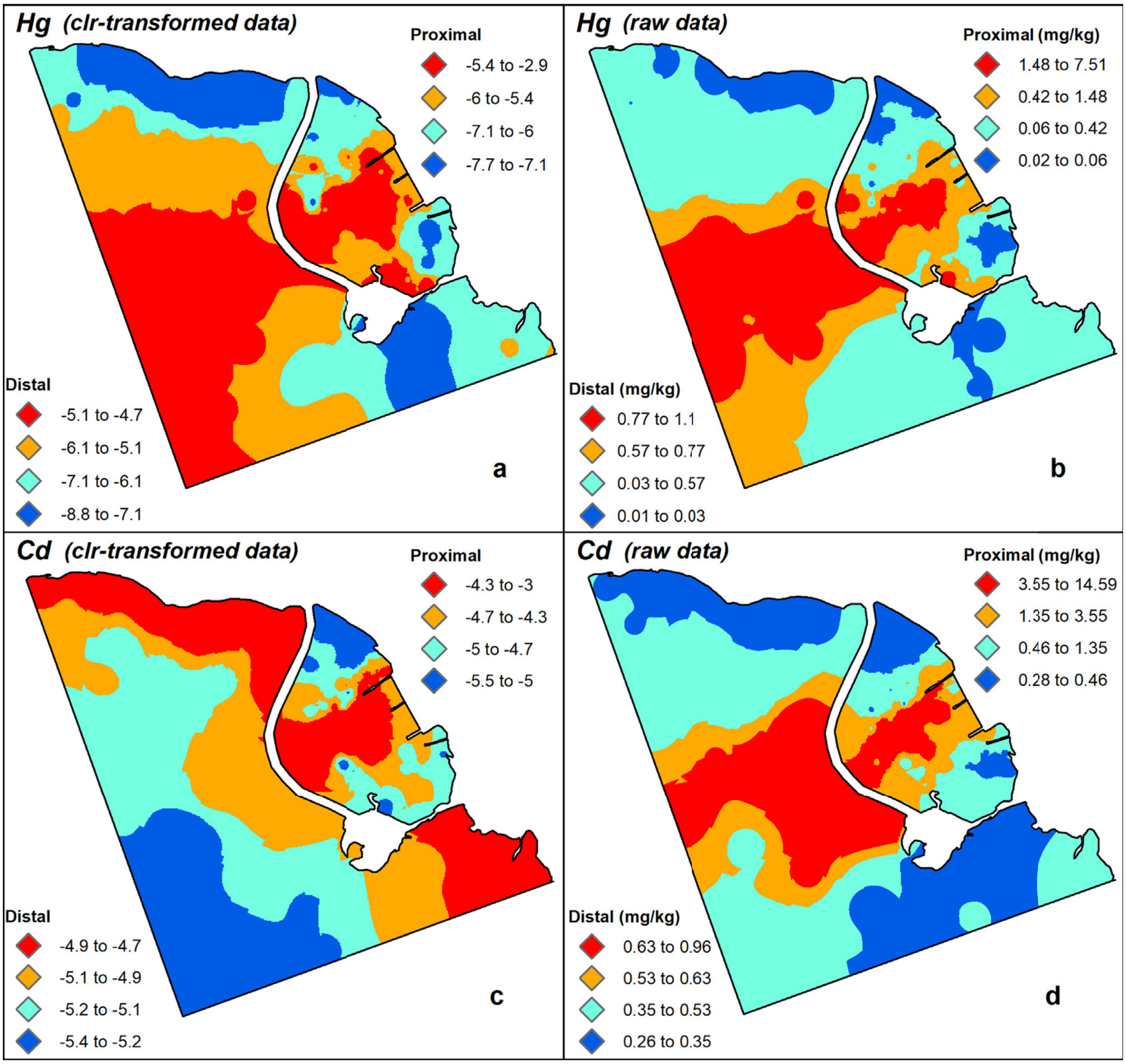


Fig. 4. Mercury and cadmium geochemical maps produced by clr-transformed data (a and c) and raw data (b and d). Opening the data does not impact spatial patterns of the PTEs, except for Cd in distal zone.

suggest a non-industrial source for Cd (e.g., from seepages), but further investigation is needed to confirm this assumption.

The elemental association of Hg, Cd, Cu, Pb and Zn in the PZN is probably related to the coastal landfill. Occurrence of OM and the fine fraction of sediments with these PTEs represents their tendency to adsorb metals in aqueous environment, being in agreement with Keshavarzi et al. (2015) and Sciarrillo et al. (2020). Moreover, the strong correlation between total polycyclic aromatic hydrocarbons (PAHs) and total organic carbon (Arienzo et al., 2017) together with similar spatial distribution of PAHs, Pb and Zn (Romano et al., 2009) in the study area support attributing this principal component to anthropogenic activity. Regarding Fig. 6a, there is a positive correlation between Hg, Cd, Cu, Pb, Zn and the distance to coastline and seepages. It might occur because of wave turbulence at coast (Arienzo et al., 2019), higher mud and OM contents far from the shoreline and insignificant impact of the seepages on the PTEs concentrations. This group of chemicals are uncorrelated with the distance to fumaroles which highlights their anthropogenic source. The similar statistical relationships between the above-mentioned variables in the DZN (Fig. 6b) feasibly indicate that the sediment pollution is extended to the deeper parts of Pozzuoli Bay. Increasing distance from fumaroles leads to higher levels of Hg and Cu in distal sediments, representing another piece of evidence for their non-geogenic origin.

The previous investigations in soil media showed that the association of Al, As and V can be related to the pyroclastic deposits of Campi Flegrei (Thiombane et al., 2019). Significant variance and correlation of these variables in PCA may indicate that As is incorporated in the crystalline structure of the sand fraction in bed sediments of PZN. Furthermore, although arsenic is uncorrelated with distance from fumaroles, it is negatively correlated with distance from coastline and seepages (Fig. 6a). Thus, increasing the distance from shoreline and seepages results in decreasing As concentration in the sediments. The present dataset does not show the relative importance of submarine arsenic-rich groundwater discharge (Aiuppa et al., 2006; De Vivo et al., 2008) in each zone. If it affects PZN and DZN equally, it can be inferred that surface discharge of mixed thermal fluids and municipal wastewaters (e.g.,

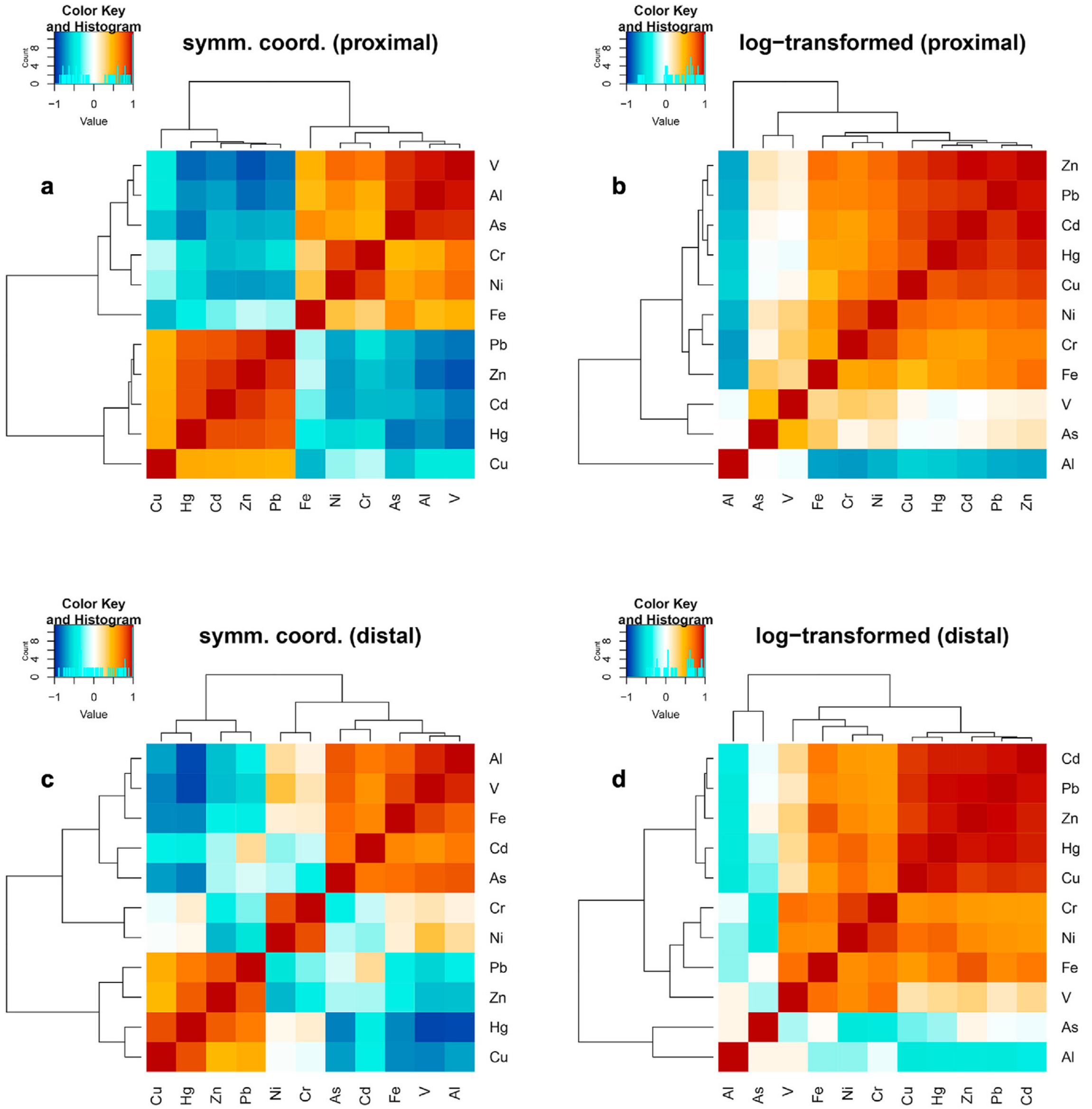


Fig. 5. The heatmaps of Spearman correlation between the measured chemicals in proximal (a and b) and distal (c and d) sediments. The left and right panels refer to the correlation analysis applied to the symmetric coordinates and log-transformed data, respectively.

the Conca d'Agnano seepage in Figs. 1 and S1) is likely the key mechanisms for transporting As to PZN. It is worth mentioning that arsenic levels in PZN are higher than DZN (Table S2 and Fig. 3b). The correlation between As, Fe and sand (Fig. 6a) may indicate that sand fraction is a potential arsenic carrier in the study area. The naturally available iron-rich sand can be a promising As adsorbent in seawater pH (Panthi and Wareham, 2011). Armiento et al. (2020) also introduced Fe oxyhydroxides as a controlling factor for As distribution in the deep sediment layers of east Pozzuoli Bay. High dependence of Ni and Fe to the sandy sediments in the proximal area is probably derived from the anthropogenic sands composed of tar and oil residues, blast furnace, slag and iron pellets (Albanese et al., 2010). In the DZN, low contribution of Al, As, V, Fe, Ni and Cr in PCA (Fig. 6b) might explain the important role of discharging thermal fluids through PZN seepages (Giglioli et al., 2020) in governing sediment geochemistry.

In order to better interpret the results of principal component analysis, the interpolated PC1 scores were represented in Fig. 7. In sediments of the proximal zone, the negative PC1 scores (i.e. the dependence of Hg, Cd, Cu, Zn and Pb to the fine fraction and OM) coincide with the Bagnoli industrial site. Presence of these polluted sediments in the offshore sector of DZN can also be verified by the Hg−Cu−mud−OM association. On the other hand, Al, As, V and Fe (together with Ni and Cr in PZN) form an association along the shoreline which might be regulated by mobilization of the fine and very fine sand that chemically interacted with seepages (Fig. 7). Hence, regarding the spatial distribution of PC1 scores and sediment dispersal towards the open sea, geochemical characteristics of the sediments in DZN likely depends on the sediments transported from PZN.

4.1.5. Sediment pollution

Considering the identified geochemical associations in robust PC1 (Fig. 6), the variables of each compositional dataset were divided in two groups (i.e. the numerators and denominators of Eqs. (6) and (7)) and the following balances were calculated for PZN

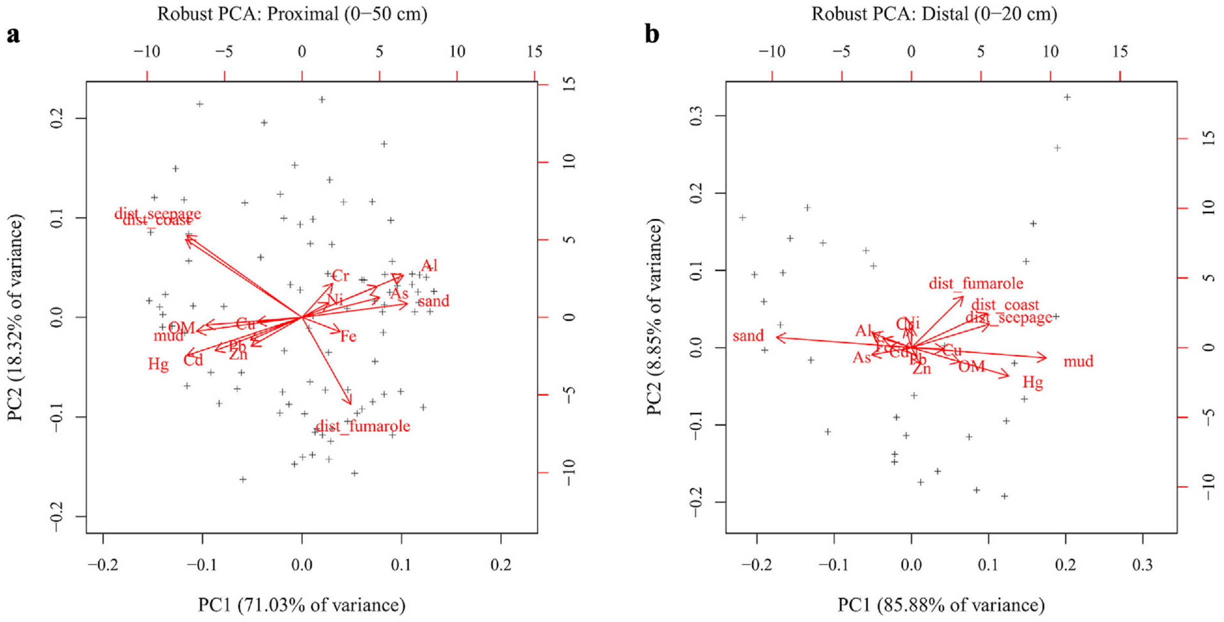


Fig. 6. Robust principal component analysis of the proximal and distal (a and b, respectively) datasets. Amongst the variables, the granulometry and element contents are two different compositions, and the rest are non-compositional variables. The distance from coast, seepage and fumarole was represented by dist_coast, dist_seepage and dist_fumarole, respectively.

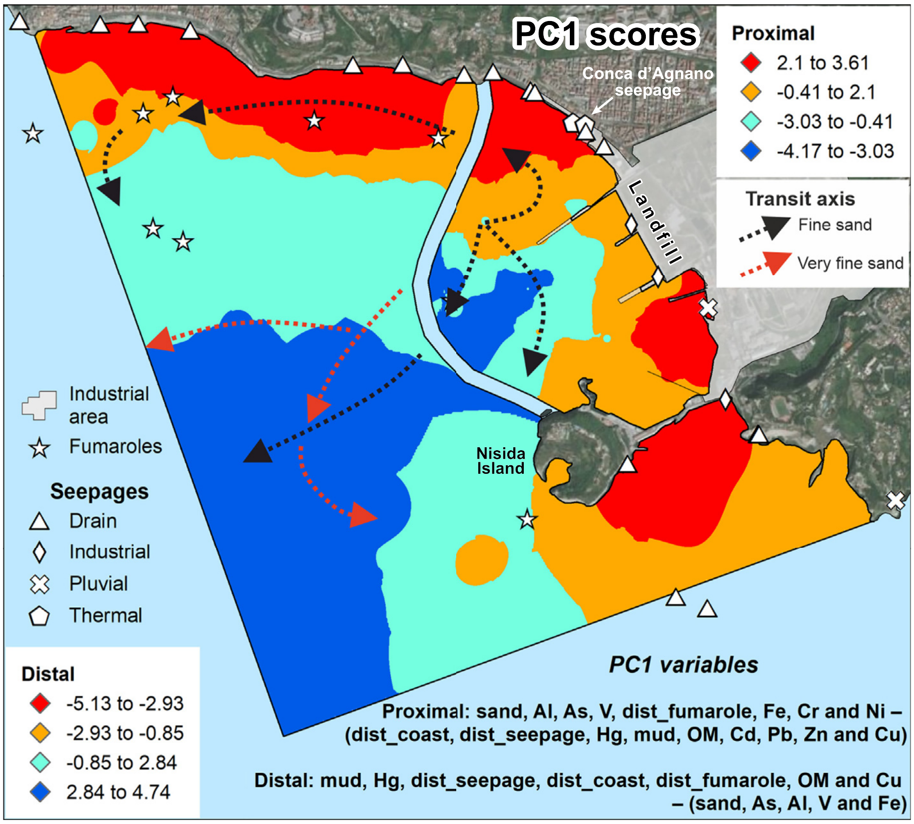


Fig. 7. PC1 score map of the sediments collected from the proximal (0–50 cm) and distal (0–20 cm) zones. Due to low contribution of Cd in robust PCA and very low correlation of Ni, Cr, Pb and Zn with PC1, they were excluded from PC1 of the DZN. The transit axes of fine and very fine sand (Arienzo et al., 2017) are also illustrated.

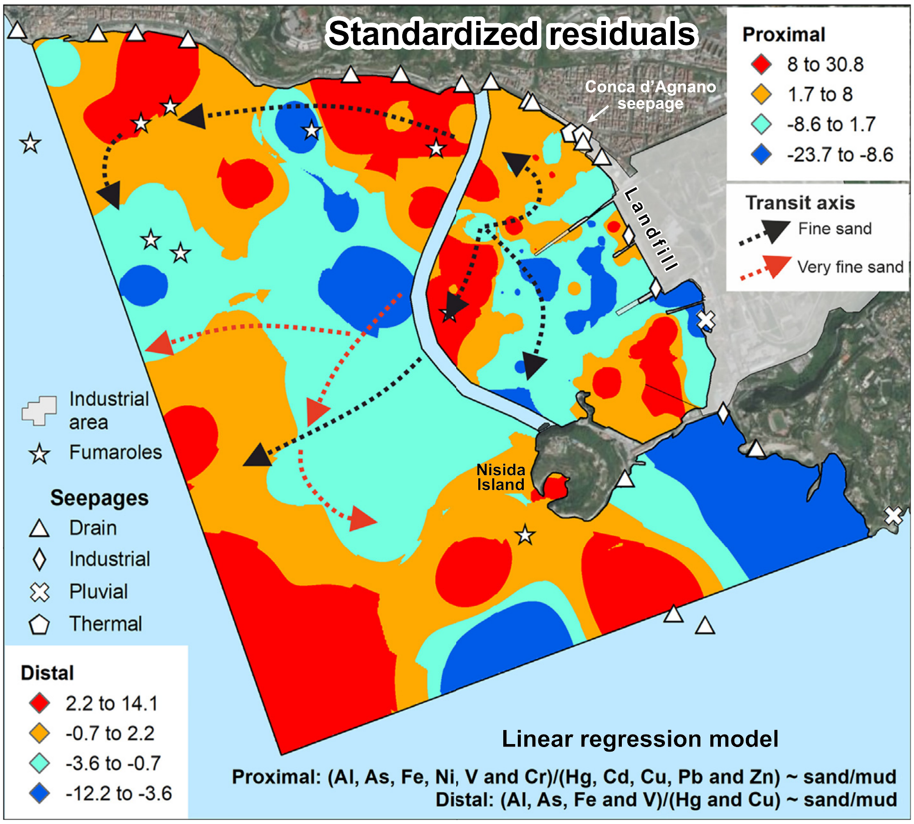


Fig. 8. Spatial distribution of the standardized residuals in the proximal and distal zones. The transit axes of fine and very fine sand (Arienzo et al., 2017) are depicted as well.

according to Eq. (2) in Text S1:

$$\tilde{z}_{\text{granulometry}} = \sqrt{\frac{1}{2}} \ln \frac{\text{sand}}{\text{mud}} \tag{6}$$

$$\tilde{z}_{\text{elements}} = \sqrt{\frac{30}{11}} \ln \frac{(\text{Al} \times \text{As} \times \text{Fe} \times \text{Ni} \times \text{V} \times \text{Cr})^{\frac{1}{6}}}{(\text{Hg} \times \text{Cd} \times \text{Cu} \times \text{Pb} \times \text{Zn})^{\frac{1}{5}}} \tag{7}$$

Likewise, the computed balances for the DZN are as follows:

$$\tilde{z}_{\text{granulometry}} = \sqrt{\frac{1}{2}} \ln \frac{\text{sand}}{\text{mud}} \tag{8}$$

$$\tilde{z}_{\text{elements}} = \sqrt{\frac{8}{6}} \ln \frac{(\text{Al} \times \text{As} \times \text{Fe} \times \text{V})^{\frac{1}{4}}}{(\text{Hg} \times \text{Cu})^{\frac{1}{2}}} \tag{9}$$

In each zone, linear regression model was then applied to the $\tilde{z}_{\text{granulometry}}$ as independent variable and $\tilde{z}_{\text{elements}}$ as dependent variable, and the obtained standardized residuals were mapped via MIDW approach as described in the section 3.2.1. As Fig. 8 shows, distribution of the positive values mostly represents the non-industrial source (i.e. pyroclastic deposits, discharge of mixed thermal fluids and municipal wastewaters, and arsenic-rich groundwater discharge) of chemicals in the sand fraction, particularly along the coastline. The high values towards the open sea likely show elevated Cr and Ni concentrations in the fine fraction (Wang et al., 2015). However, the samples with negative standardized residuals occur in front of Bagnoli industrial site which can also be traced in the DZN. The negative values in south and east of DZN might be attributed to the seepages, especially the industrial seepage. Wave-induced sediment transport probably disturbed the spatial patterns.

4.2. Sediment geochemistry along the cores

Considering the sediments from different depths (Tables S2 and S3) in PZN, the average OM contents and CVR values are higher at surface (<100 cm depth). The same vertical pattern was also observed for mud, Cu, Pb, Zn, Cd, Hg, Cr, Ni and Fe. On the other hand, variation of sand, Al, As and V are almost similar. In the next step, the sediment cores that reach 300 cm depth (the bottom right panel of Fig. 9) were selected and the average values for the analyzed elements, grain size fractions and OM in each depth interval were plotted in Fig. 9. Furthermore, the clr-transformed values of the metal(loid)s and granulometry data were included in each graph. It is evident that opening the data changes the vertical trend of some parameters (e.g., Ni and Cr) and better illustrates variation of As and V with depth. Regarding the clr-transformed values (except for OM), two groups of variables can be identified: (1) Al, As, Cr, V, Ni and sand; and (2) Cu, Cd, Hg, Pb, Zn, mud and OM. The minerals such as feldspar (plagioclase and potassium feldspar), biotite and pyroxene in the sand fraction (Romano et al., 2018) can be source of the elements in the first group. It seems that Fe shows an increasing trend towards depth, but high Fe contents at surface might be due to the iron pellets in the sea sediments (Albanese et al., 2010). In addition, the high proportion of haematite in the 0–50 cm depth interval of the sediment cores collected between the piers (Molisso et al., 2020) and the coarse fragments of coal and blast furnace slag in the upper 120 cm layer of sediments (Romano et al., 2018) can lead to the unregular Fe profile. Vertical profiles of the elements in the second group likely suggest that although industrial activity was stopped several decades ago, Bagnoli landfill is still a pollution source. Ignoring compositional nature of the geochemical data avoids identifying some of the observed patterns in Fig. 9.

Because clr-transformation is sub-compositionally incoherent, PCA was also performed for dimension reduction and for verifying the above-mentioned visual inspections. The variation retained by the first two principal components of robust PCA is approximately 6% greater than that of classical PCA (93.43% in Fig. 10 vs. 87.19% in Fig. S12). Therefore, biplot of robust PC1 and PC2 is more informative. Two associations of parameters can be determined in Fig. 10a: (1) Hg, Cd, Cu, Pb, Zn, mud and OM; and (2) Al, As, V, Fe, Ni, Cr and sand. This classification of elements is consistent with the observed trends in Fig. 9. There is a strong association between Hg, Cd, Cu, Pb, Zn, mud and OM at <100 cm depth. These PTEs are anthropogenic pollutants as it was discussed previously. However, Al, As, V, Fe, Ni, Cr and sand are highly correlated in deeper parts (100–300 cm) which probably represents their geogenic source. Armiento et al. (2020) attributed the elevated arsenic contents in the sediments from 200 to 300 cm below surface to iron oxy-hydroxides. The same groups of variables in PCA of the sediments from 0 to 50 cm depth (Fig. 6a) and the interpolated map of PC1 scores (Fig. 7) indicates extent of the polluted sediments in front of Bagnoli coastal landfill. According to Fig. 10, thickness of this polluted layer is about 100 cm along the shoreline. Sharp and Nardi (1987) estimated a sedimentation rate of 0.4 cm/y in the study area and concluded that maximum pollution occurred approximately 70 years ago. Because the current sampling campaign was conducted in 2017, it can be inferred that the most polluted sediment layer is about 40 cm below surface. However, some mechanisms such as sediment burial (Daliri et al., 2020) and wave turbulence (Arienzo et al., 2019) might affect vertical distribution of chemicals at <100 cm depth. Diagenesis is another process that might alter chemical phases and geochemical behaviors of elements, but further investigation is required to confirm this assumption. In the form biplot (Fig. 10b) which favors visualization of individuals, the existing trend from natural samples at depth to contaminated samples at surface depicts pollution history of the study area.

The discussion about subsurface geochemistry of marine sediments can be informative in the present study because the profiles of variables (e.g., Cu and Fe in Figs. S13 and S14, respectively) generally agree with the high-resolution measurements along

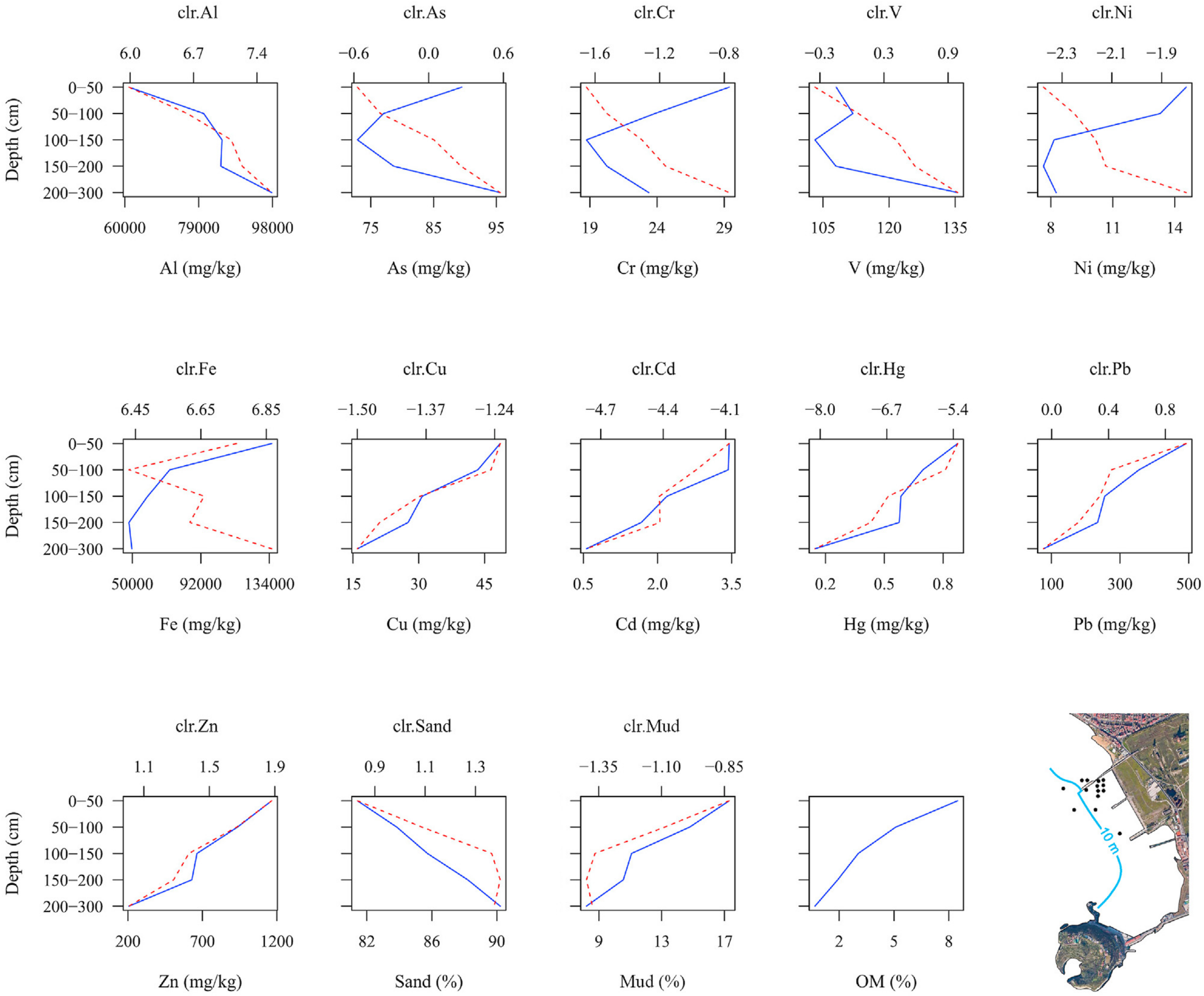


Fig. 9. Variation of the chemical elements, granulometry and organic matter (OM) content of the sediments from 0 to 300 cm depth. The solid blue line represents the measured values and the dashed red line shows the clr-transformed data. Location of the selected core samples are indicated in the bottom right map. For interpretation of the references to the colors in this figure, the reader is referred to the Web version of this article.

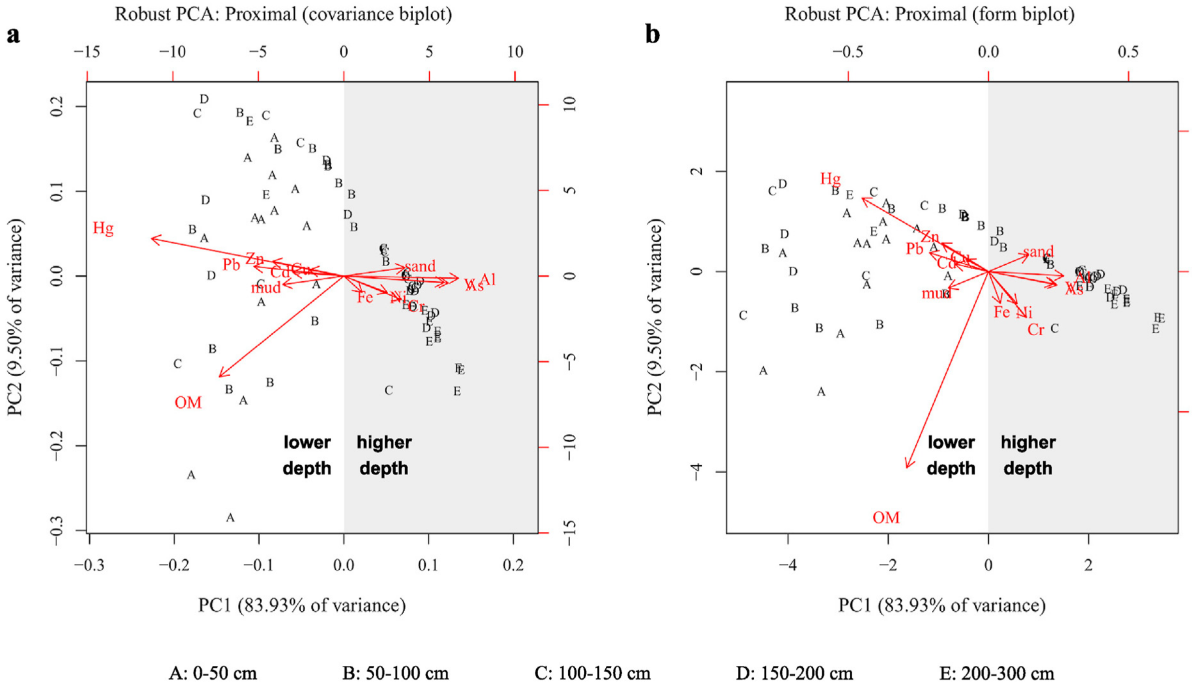


Fig. 10. Robust principal component analysis of the variables analyzed along the selected cores (the bottom right panel of Fig. 9) in the proximal zone: a) covariance biplot and b) form biplot. OM is a non-compositional variable whilst the granulometry and elements are two different compositions.

isolated cores (Damiani et al., 1987; Romano et al., 2018). Nevertheless, the site-dependent geochemical signatures (e.g., Figs. 10, S13 and S14) in front of the industrial plant highlight complexity of the sedimentary environment and importance of number and length of sediment cores together with sampling resolution along each core.

5. Conclusions and future work

The existing literature indicates that geochemistry of the sediments deposited in the eastern sector of Pozzuoli Bay has been affected by natural (e.g., underwater fumaroles, minerals and submarine groundwater discharge) and anthropogenic (for instance, industrial activity in the past, the ineffective coastal landfill and the seepages) factors. Researchers proposed different sources for PTEs (particularly for As) depending on the variables in the dataset and data treatment techniques. In the present investigation, the study area is divided in the proximal and distal zones (PZN and DZN, respectively). In geochemical maps and along the sediment cores, the clr-transformed values of metal(loid)s evidence meaningful univariate patterns. According to the robust PCA of PZN, two strong associations are detected: (1) Hg, Cd, Cu, Pb, Zn, OM, mud, dist_coast and dist_seepage; and (2) Al, As, V, Fe, Cr, Ni, sand and dist_fumarole. The former can be attributed to the industrial activity because concentrations of the PTEs decease along the cores and in the DZN, but the latter is likely contributed by pyroclastic deposits of Campi Flegrei and the seepages with hydrothermal component. It seems that submarine groundwater discharge, underwater fumaroles and the coastal landfill do not play the major role in arsenic geochemistry. Despite almost similar associations in DZN, low contribution of variables in the robust PCA probably suggests that the main sources of elements are spatially located in the proximal zone. Therefore, this pioneering work demonstrates the effectiveness of compositional data analysis for bridging the ongoing knowledge gaps in sediment geochemistry of complex coastal environments.

The following future research directions are suggested to shed light on the geochemistry of sea sediments in Pozzuoli Bay: (1) applying the compositional and fractal approaches to the seafloor surface sediments of the same depth; (2) using CoDA technique after collecting spatially dense cores and obtaining high-resolution measurements along them to reduce uncertainty; (3) considering a buffer zone around the Bagnoli-Coroglio site of national interest (SIN), according to the sediment movement path, for the future monitoring plans; (4) investigating chemical forms of PTEs (e.g., the fraction of metals bound in carbonates, Fe–Mn oxides, sulfides, organic matter and lithogenic minerals) in order to define chemical interaction, mobility, bioavailability and potential toxicity of PTEs; and (5) extending the ideas in this article to other pollutants such as PAHs.

Author contributions

Renato Somma: Conceptualization, Investigation, Writing – review & editing, Supervision, Project administration; Pooria Ebrahimi: Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing; Claudia Troise: Writing – review & editing; Giuseppe De Natale: Writing – review & editing; Annalise Guarino: Formal analysis, Visualization; Domenico Cicchella: Writing – review & editing; Stefano Albanese: Writing – review & editing, Supervision, Project administration

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to express their gratitude to Antonio Aruta from Department of Earth, Environmental and Resources

Sciences at the University of Naples Federico II (Naples, Italy) and Salvatore Dominech from School of Ocean and Earth Science at Tongji University (Shanghai, China) for their support in multifractal inverse distance weighted (MIDW) mapping of clr-transformed values. In addition, we acknowledge Zoological Station 'Anton Dohrn', Institute for the Environmental Protection and Research (ISPRA), National Research Center (CNR) and INVITALIA for providing the metal(loid)s dataset and the supplementary information about Pozzuoli Bay. The thanks are also extended to Giovanna Armiento from Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) for providing the information on accuracy of the analytical results. Further, the anonymous reviewers are appreciated for their constructive comments and useful suggestions which helped to improve the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemosphere.2021.129955.

Funding

This study was mainly supported by the ABBaCo project [grant number C62F16000170001] and partially supported by the PON-S4E [grant number cod. SCN 00393] and PON-OT4CLIMA [grant number cod. ARS01 00405] projects, all funded by the Italian Ministry for Education, University and Research.

References

ABBaCo Project, 2018. Sperimentazioni Pilota Finalizzate al Restauro Ambientale e Balneabilità del SIN Bagnoli-Coroglio. Available online: http://www.szn.it/index.php/en/research/integrative-marine-ecology/research-projects-emi/abbaco. (Accessed 28 October 2020).

Aitchison, J., 1982. The statistical analysis of compositional data. J. Roy. Stat. Soc. B Met. 44 (2), 139–160.

Aitchison, J., Greenacre, M., 2002. Biplots of compositional data. J. Roy. Stat. Soc. C-App. 51 (4), 375–392.

Aiuppa, A., Avino, R., Brusca, L., Caliro, S., Chiodini, G., D'Alessandro, W., Favara, R., Federico, C., Ginevra, W., Inguaggiato, S., Longo, 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). Chem. Geol. 229 (4), 313–330.

Albanese, S., De Vivo, B., Lima, A., Cicchella, D., Civitillo, D., Cosenza, A., 2010. Geochemical baselines and risk assessment of the Bagnoli brownfield site coastal sea sediments (Naples, Italy). J. Geochem. Explor. 105 (1–2), 19–33.

Álvarez-Vázquez, M.Á., Hošek, M., Elznicová, J., Pacina, J., Hron, K., Fačevicová, K., Talská, R., Bábek, O., Grygar, T.M., 2020. Separation of geochemical signals in fluvial sediments: new approaches to grain-size control and anthropogenic contamination. Appl. Geochem. 123, 104791.

Arienzo, M., Donadio, C., Mangoni, O., Bolinesi, F., Stanislao, C., Trifuoggi, M., Toscanesi, M., Di Natale, G., Ferrara, L., 2017. Characterization and source apportionment of polycyclic aromatic hydrocarbons (PAHs) in the sediments of gulf of Pozzuoli (Campania, Italy). Mar. Pollut. Bull. 124 (1), 480–487.

Arienzo, M., Toscanesi, M., Trifuoggi, M., Ferrara, L., Stanislao, C., Donadio, C., Villari, G., De Vico, G., Carella, F., 2019. Contaminants bioaccumulation and pathological assessment in Mytilus galloprovincialis in coastal waters facing the brownfield site of Bagnoli. Italy. Mar. Pollut. Bull. 140, 341–352.

Armiento, G., Caprioli, R., Cerbone, A., Chiavarini, S., Crovato, C., De Cassan, M., De Rosa, L., Montereali, M.R., Nardi, E., Nardi, L., Pezza, M., 2020. Current status of coastal sediments contamination in the former industrial area of Bagnoli-Coroglio (Naples, Italy). Chem. Ecol. 1–19.

Beck, K.K., Mariani, M., Fletcher, M.S., Schneider, L., Aquino-López, M.A., Gadd, P.S., Heijnis, H., Saunders, K.M., Zawadzki, A., 2020. The impacts of intensive mining on terrestrial and aquatic ecosystems: a case of sediment pollution and calcium decline in cool temperate Tasmania, Australia. Environ. Pollut. 265, 114695.

Bergamin, L., Romano, E., Magno, M.C., Ausili, A., Gabellini, M., 2005. Pollution monitoring of bagnoli Bay (Tyrrhenian Sea, Naples, Italy), a sedimentological, chemical and ecological approach. Aquat. Ecosyst. Health 8 (3), 293–302.

Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (Eds.), 2006. Compositional Data Analysis in the Geosciences: from Theory to Practice. Special Publications, London, p. 264. Geological Society.

Daliri, M., Somma, R., Buccino, M., Troise, C., Molisso, F., De Natale, G., 2020. Numerical analysis of arsenic concentration in the marine sediments of bagnoli Bay. In: The 30th International Ocean and Polar Engineering Conference. International Society of Offshore and Polar Engineers, 2020, September.

Damiani, V., Baudo, R., De Rosa, S., De Simone, R., Ferretti, O., Izzo, G., Serena, F., 1987. A case study: Bay of Pozzuoli (gulf of Naples, Italy). In: Thomas, R.L., Evans, R., Hamilton, A.L., Munawar, M., Reynoldson, T.B., Sadar, M.H. (Eds.), Ecological Effects of in Situ Sediment Contaminants, Developments in Hydrobiology, vol. 39. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-4053-6_21.

De Natale, G., Troise, C., Mark, D., Mormone, A., Piochi, M., Di Vito, M.A., Isaia, R., Carlino, S., Barra, D., Somma, R., 2016. The Campi Flegrei deep drilling project (CFDDP): new insight on caldera structure, evolution and hazard implications for the Naples area (southern Italy). G-cubed 17 (12), 4836–4847.

De Pippo, T., Donadio, C., Pennetta, M., Terlizzi, F., Vecchione, C., Vegliante, M., 2002. Seabed morphology and pollution along the Bagnoli coast (Naples, Italy): a hypothesis for environmental restoration. Mar. Ecol. 23, 154–168.

De Vivo, B., Belkin, H., Lima, A., 2008. Environmental Geochemistry: Site Characterization, Data Analysis and Case Histories. Elsevier, Oxford, UK.

Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S., Albanese, S., 2020a. Potentially Toxic Elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (Southern Italy): a focus on cadmium. Proceedings of I International Meeting of Geohealth Scientists 35–40. https://doi.org/10.14644/ghc2020.006.

Dominech, S., Guarino, A., Aruta, A., Ebrahimi, P., Yang, S., Albanese, S., 2020b. Potentially toxic elements (PTEs) distribution and main geochemical processes in Sabato river catchment basin (southern Italy). In: Fiore, S. (Ed.), I International Meeting of GeoHealth Scientists GHC 2020 BARI. Italy. Digilabs Pub., Bari, Italy, p. 51.

Ebrahimi, P., 2013. Environmental Geochemistry and Bioavailability of Heavy Metals in Coastal and Marine Sediments of Chabahar Bay, SE Iran. Master's thesis. Shiraz University, Shiraz, Iran. https://doi.org/10.13140/RG.2.2.17519.66722.

Egozcue, J.J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Math. Geol. 37 (7), 795–828.

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

Filzmoser, P., 1999. Robust principal component and factor analysis in the geostatistical treatment of environmental data. Environmetrics 10 (4), 363–375.

Filzmoser, P., Hron, K., Reimann, C., 2009. Principal component analysis for compositional data with outliers. Environmetrics 20 (6), 621–632.

Filzmoser, P., Hron, K., Templ, M., 2018. Applied Compositional Data Analysis. With Worked Examples in R. Springer Nature, Switzerland. https://doi.org/10.1007/978-3-319-96422-5.

Flood, R.P., Bloemsma, M.R., Weltje, G.J., Barr, I.D., O'Rourke, S.M., Turner, J.N., Orford, J.D., 2016. Compositional data analysis of Holocene sediments from the West Bengal Sundarbans, India: geochemical proxies for grain-size variability in a delta environment. Appl. Geochem. 75, 222–235.

Giglioli, S., Colombo, L., Contestabile, P., Musco, L., Armiento, G., Somma, R., Vicinanza, D., Azzellino, A., 2020. Source apportionment assessment of marine sediment contamination in a post-industrial area (Bagnoli, Naples). Water 12 (8), 2181.

ICRAM, 2005. Caratterizzazione ambientale dei fondali e degli arenili inclusi nella perimetrazione del sito di bonifica di Napoli Bagnoli-Coroglio. Relazione tecnica conclusiva. BoI-Pr-CA-BA-relazione-02.04. Il Commissario di Governo per le Bonifiche e Tutela delle Acque nella regione Campania delegato ex O.P.C.M. 2425/96 e successive, p. 113.

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. Chem. Erde-Geochem. 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. Math. Geosci. 49 (6), 777–796.

Liu, Y., Cheng, Q., Zhou, K., Xia, Q., Wang, X., 2016. Multivariate analysis for geochemical process identification using stream sediment geochemical data: a perspective from compositional data. Geochem. J. 50 (4), 293–314.

Liu, Y., Zhou, K., Carranza, E.J.M., 2018. Compositional balance analysis for geochemical pattern recognition and anomaly mapping in the western Junggar region, China. Geochem. Explor. Environ. Anal. 18 (3), 263–276.

Lloyd, C.D., Pawlowsky-Glahn, V., Egozcue, J.J., 2012. Compositional data analysis in population studies. Ann. Assoc. Am. Geogr. 102 (6), 1251–1266.

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E.L., Anna di Palma, M., 2020. Robustbase: Basic Robust Statistics. R Package Version 0, pp. 93–96. http://robustbase.r-forge.r-project.org/.

Maronna, R., Martin, R.D., Yohai, V.J., 2006. Robust Statistics: Theory and Methods. John Wiley, New York.

Molisso, F., Caccavale, M., Capodanno, M., Di Gregorio, C., Gilardi, M., Guarino, A., Oliveri, E., Tamburrino, S., Sacchi, M., 2020. Sedimentological analysis of marine deposits off the bagnoli-coroglio site of national interest (SNI), Pozzuoli (Napoli) Bay. Chem. Ecol. 36 (6), 565–578.

Morroni, L., d'Errico, G., Sacchi, M., Molisso, F., Armiento, G., Chiavarini, S., Rimauro, J., Guida, M., Siciliano, A., Ceparano, M., Aliberti, F., 2020. Integrated characterization and risk management of marine sediments: the case study of the industrialized Bagnoli area (Naples, Italy). Mar. Environ. Res. 104984.

Panthi, S.R., Wareham, D.G., 2011. Removal of arsenic from water using the adsorbent: New Zealand iron-sand. J. Environ. Sci. Heal. A 46 (13), 1533–1538.

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. Proc. Roy. Soc. Lond. 60 (359–367), 489–498.

Petrik, A., Thiombane, M., Albanese, S., Lima, A., De Vivo, B., 2018. Source patterns of Zn, Pb, Cr and Ni potentially toxic elements (PTEs) through a compositional discrimination analysis: a case study on the Campanian topsoil data. Geoderma 331, 87–99.

Razum, I., Miko, S., Ilijanić, N., Hasan, O., Miko, M.Š., Brunović, D., Pawlowsky-Glahn, V., 2020. A compositional approach to the reconstruction of geochemical processes involved in the evolution of Holocene marine flooded coastal karst basins (Mljet Island, Croatia). Appl. Geochem. 116, 104574.

Reimann, C., Filzmoser, P., Garrett, R., Dutter, R., 2008. Statistical Data Analysis Explained: Applied Environmental Statistics with R. John Wiley & Sons. https://doi.org/10.1002/9780470987605.

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. Sci. Total Environ. 607, 965–971.

Rolandi, G., De Natale, G., Kilburn, C.R., Troise, C., Somma, R., Di Lascio, M., Fedele, A., Rolandi, R., 2020. The 39 ka Campanian Ignimbrite eruption: new data on source area in the Campanian Plain. Vesuvius, Campi Flegrei, and Campanian Volcanism. Elsevier, pp. 175–205. https://doi.org/10.1016/b978-0-12-816454-9.00008-0.

Romano, E., Ausili, A., Zharova, N., Magno, M.C., Pavoni, B., Gabellini, M., 2004. Marine sediment contamination of an industrial site at port of bagnoli, gulf of Naples. Southern Italy. Mar. Pollut. Bull. 49 (5–6), 487–495.

Romano, E., Bergamin, L., Ausili, A., Pierfranceschi, G., Maggi, C., Sesta, G., Gabellini, M., 2009. The impact of the Bagnoli industrial site (Naples, Italy) on sea-bottom environment. Chemical and textural features of sediments and the related response of benthic foraminifera. Mar. Pollut. Bull. 59 (8–12), 245–256.

Romano, E., Bergamin, L., Magno, M.C., Pierfranceschi, G., Ausili, A., 2018. Temporal changes of metal and trace element contamination in marine sediments due to a steel plant: the case study of Bagnoli (Naples, Italy). Appl. Geochem. 88, 85–94.

Romano, E., Gabellini, M., 2001. Analisi delle caratteristiche granulometriche – sedimenti, scheda 3. In: Cicero, A.M., Di Girolamo, I. (Eds.), Metodologie Analitiche di Riferimento del Programma di Monitoraggio per il controllo dell'ambiente marino costiero (triennio 2001–2003). Ministero dell'Ambiente e della Tutela del Territorio, ICRAM, Roma, p. 122.

Schlee, J.S., 1973. Atlantic Continental Shelf and Slope of the United States Sediment Texture of the Northeastern Part. U.S. Geological Survey Professional Paper 529-L, p. 64. https://doi.org/10.3133/pp529L.

Sciarrillo, R., Zuzolo, D., Cicchella, D., Iannone, F., Cammino, G., Guarino, C., 2020. Contamination and ecological risk assessment of the seaport of Naples (Italy): insights from marine sediments. J. Geochem. Explor. 210, 106449.

Selinus, O., 2013. Essentials of Medical Geology: Revised Edition. Springer Netherlands, p. 805. https://doi.org/10.1007/978-94-007-4375-5.

Sharp, W.E., Nardi, G., 1987. A study of the heavy-metal pollution in the bottom sediments at Porto Di Bagnoli (Naples), Italy. J. Geochem. Explor. 29 (1–3), 31–48.

Somma, R., Iuliano, S., Matano, F., Molisso, F., Passaro, S., Sacchi, M., Troise, C., De Natale, G., 2016. High-resolution morpho-bathymetry of Pozzuoli Bay, southern Italy. J. Maps 12 (2), 222–230. https://doi.org/10.1080/17445647.2014.1001800.

Thiombane, M., Di Bonito, M., Albanese, S., Zuzolo, D., Lima, A., De Vivo, B., 2019. Geogenic versus anthropogenic behaviour and geochemical footprint of Al, Na, K and P in the Campania region (Southern Italy) soils through compositional data analysis and enrichment factor. Geoderma 335, 12–26.

Thiombane, M., Zuzolo, D., Cicchella, D., Albanese, S., Lima, A., Cavaliere, M., De Vivo, B., 2018. Soil geochemical follow-up in the Cilento World Heritage Park (Campania, Italy) through exploratory compositional data analysis and CA fractal model. J. Geochem. Explor. 189, 85–99.

Trifuoggi, M., Donadio, C., Mangoni, O., Ferrara, L., Bolinesi, F., Nastro, R.A., Stanislao, C., Toscanesi, M., Di Natale, G., Arienzo, M., 2017. Distribution and enrichment of trace metals in surface marine sediments in the Gulf of Pozzuoli and off the coast of the brownfield metallurgical site of Ilva of Bagnoli (Campania, Italy). Mar. Pollut. Bull. 124 (1), 502–511.

Troise, C., De Natale, G., Schiavone, R., Somma, R., Moretti, R., 2019. The Campi Flegrei caldera unrest: discriminating magma intrusions from hydrothermal effects and implications for possible evolution. Earth Sci. Rev. 188, 108–122.

Wang, M., Albanese, S., Lima, A., Cannatelli, C., Cosenza, A., Lu, W., Sacchi, M., Doherty, A., De Vivo, B., 2015. Compositional analysis and pollution impact assessment: a case study in the Gulfs of Naples and Salerno. Estuar. Coast Shelf Sci. 160, 22–32.

Wang, W., Zhao, J., Cheng, Q., 2014. Mapping of Fe mineralization-associated geochemical signatures using logratio transformed stream sediment geochemical data in eastern Tianshan, China. J. Geochem. Explor. 141, 6–14.

Publication No. 4

Zheng, C., Luo, X., Wen, M., Ebrahimi, P., Liu, P., Liu, G., ... & Albanese, S. (2020). Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China. Journal of Geochemical Exploration, 213, 106509. https://doi.org/10.1016/j.gexplo.2020.106509.

Journal of Geochemical Exploration 213 (2020) 106509

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

journal homepage: www.elsevier.com/locate/gexplo

Axial primary halo characterization and deep orebody prediction in the Ashele copper-zinc deposit, Xinjiang, NW China

Chaojie Zheng[a,b], Xianrong Luo[a,*], Meilan Wen[a,*], Pooria Ebrahimi[b], Panfeng Liu[a,c], Gang Liu[a], Heping Li[d], Chaocheng Wei[d], Zesu Chen[d], Stefano Albanese[b]

[a] Guilin University of Technology, College of Earth Sciences, 319 Yanshan Street, Yanshan, Guilin 541006, China
[b] Department of Earth, Environment and Resources Sciences, University of Naples Federico II, Naples, Italy
[c] Guilin University of Technology, College of Environmental Sciences and Engineering, 319 Yanshan Street, Yanshan, Guilin 541006, China
[d] Xinjiang Habahe Asher Copper Co., Ltd. Kaba, Altay, Xinjiang, China

ARTICLE INFO

Keywords:
Ashele
Xinjiang
Primary halo zoning
Copper-zinc polymetallic deposit
R-mode factor analysis
Deep prediction

ABSTRACT

In this investigation, a total of 404 rock and ore samples were systematically collected from 5 boreholes in two exploration lines of the Ashele copper-zinc polymetallic mining area in Xinjiang. Twelve elements were analyzed to evaluate the potential for deep and periphery prospecting. Subsequently, the concentration zoning and the vertical element zoning sequence of the primary halo were graphically generated. Based on both previous studies and the result of the actual analysis, it was determined that, for Ashele mining area, As, Sb, Ba, and Au are supra-ore halo elements, Pb, Zn, Cu, and Ag (Au) are near-ore halo elements, and Bi, Co, Mo, and Sn are sub-ore halo elements. Factor analysis revealed exhalative-sedimentation mineralization as the first metallogenic stage followed by metamorphic-deformation and hydrothermal superimposed mineralization as the second one. The superposition area of the sub-ore halo and the supra-ore halo and multivariate prospecting information (indicated by the primary halo zoning index, the geochemical parameter, the trilinear diagram of the primary halo, and the factor score map) suggest that there may be blind orebodies in the deep parts of four boreholes. Some geochemical parameters revealed the multiple-periodic and multiple staged overlapping the signatures of the mineralization.

1. Introduction

Primary halo zonation was originally defined by Safronov (1936) for prospecting deep blind orebodies and predicting resource potentials (Carranza, 2012; Li et al., 2016b). Primary halos are zones surrounding metalliferous deposits and ore bodies, created by the interactions between ore-forming fluids and enclosing rocks, where several chemical elements are introduced and redistributed during the processes of ore formation (Parsapoor et al., 2017; Beus and Grigorian, 1977; Carranza, 2012; Goldberg et al., 2003; Gong et al., 2016; Wang et al., 2013; Wang et al., 2011). Different trace elements (i.e. geochemical indicators) in the multi-component halos can form different halo patterns (Wang et al., 2013). The characteristics and scale of the halo zonation are related to the concentration of elements in the ore-bearing solutions, the relative mobility of elements, the background content of elements in the enclosing rocks, and mineralization types (Beus and Grigorian, 1977). The study of a large number of hydrothermal deposits reveals that primary halos may be dozens or hundreds of times larger than orebodies they express (Parsapoor et al., 2017; Chen and Zhao, 1998; Hosseini-Dinani and Aftabi, 2016; Helba et al., 2020) leading to visible primary halos of some deeply buried deposits on the surface (Liu et al., 2010).

Three typical zonations including axial, vertical and transversal zonations are characterized in the halos of mineral deposits (Parsapoor et al., 2017; Beus and Grigorian, 1977; Wang et al., 2013). The axial zonation expresses flow direction of ore-bearing solutions, whereas the transversal zonation in primary halos reflects different mineralizations across the orebody (Beus and Grigorian, 1977; Chen and Zhao, 1998). It was supposed that, in steeply dipping orebody, the axial zonation is in accord with the vertical zonation, while in sub-horizontal orebodies, it coincides with horizontal zonation (Beus and Grigorian, 1977; Luo et al., 2007; Shao, 1997). Most research projects on the axial zoning show that this approach is highly effective for exploration of non-cropping and hydrothermal polymetallic deposits (Chen et al., 2008; Hannington et al., 2003; Ziaii et al., 2011; Wang et al., 2013; Wang et al., 1995; Wei et al., 2000; Yousefi, 2017; Zhao et al., 1989).

---

* Corresponding authors.
E-mail addresses: lxr811@glut.edu.cn (X. Luo), meilanwen112@126.com (M. Wen).

https://doi.org/10.1016/j.gexplo.2020.106509
Received 19 November 2019; Received in revised form 11 February 2020; Accepted 24 February 2020

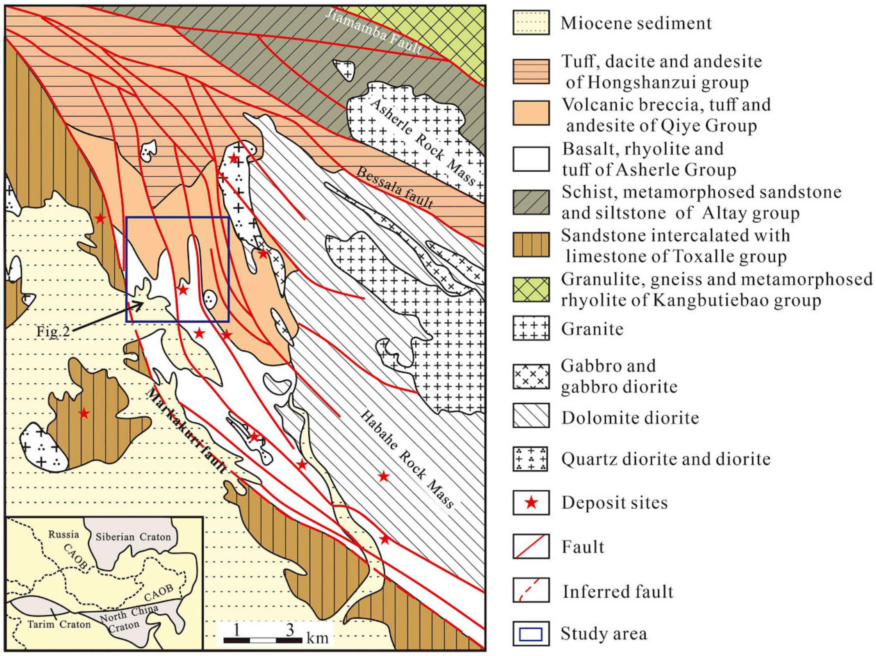


Fig. 1. Geological map of Ashele Basin.
(Adapted from Yang et al., 2016).

Various mineral exploration techniques based on primary halo have been developed within the last two decades. The orebody-scale approach is a precise method in assessing the potential for prospecting deep hydrothermal deposits compared to the regional scale and deposit scale methods. Different hydrothermal deposits show a decreasing trend in concentrations of some lithogeochemical elements with increasing distance from the orebody, explicitly indicating zonation of metallogenic elements. However, occurrence of another deep orebody is sometimes confirmed by a continuous increase in concentration of elements towards depths. Based on the above-mentioned discussion, the structurally superimposed haloes and the zoning sequence methods were developed to investigate the primary halos for deep mineral exploration (Konstantinov and Strujkov, 1995; Li et al., 2010, 2013; Li et al., 1995; Liu et al., 2014). Applying these methods facilitates prediction of mineral deposits at great depth.

The Ashele copper-zinc deposit is a typical VMS type deposit located in the Ashele basin on the southwest margin of the Altai orogenic belt in China (Fig. 1). The main metallic minerals are copper, and zinc associated with gold and silver. Previous studies, including several compositional, structural, genetic and geological aspects were carried out in detail systematically, and significant results have been achieved (Liu et al., 2010; Wu et al., 2015). However, the primary halo of the deposit has not been studied systematically. The work presented here aims to provide the basis for the prospecting prediction of the deep and peripheral sections of the mining area. To this end, the primary halo zoning theory of hydrothermal deposits was considered, and the characteristics of concentration zoning, vertical zoning, and trace element geochemistry of the primary halos were studied.

2. Geological setting

2.1. Regional geology

The Ashele copper-zinc deposit is located in the Akhler volcano-sedimentary basin on the southern margin of the Altai geosynclinal fold system, the Qiongkul-Abagong fold belt and the Altai orogenic belt (Fig. 1). The study area belongs to the late Paleozoic interarc extensional basin in the southern margin of the Central Asian orogenic belt.

The regional stratigraphic setting mainly comprises marine acidic volcanic rocks, pyroclastic rocks and terrigenous clastic rocks of the Upper Silurian-Lower Devonian Kangbutiebao group, and marine terrigenous clastic rocks intercalated with siliceous rocks and carbonate rocks of the Lower Devonian Tuoxale group. Marine flysch clastic sedimentary rocks of the Middle Devonian Altay group, intermediate-acid and basic volcanic rocks, and pyroclastic sedimentary rocks of the Middle Devonian Ashele group, are the predominant host strata of the orebody in the Ashele basin. Intermediate-basic volcanic rocks and pyroclastic sedimentary rocks are in the bathyal facies of the Upper Devonian Qiya group; Neritic facies of volcanic and clastic sedimentary rocks, intercalated with carbonate rocks, are in the Lower Carboniferous Hongshanzui group. (Fig. 1) (Yang et al., 2013; Zheng et al., 2015).

There are many large fold structures in the region: the Kuludeneng synclinal and the Jiamamba complex anticline, for example, are tight closed linear folds running roughly N—W. The faults contain the regional Markakuri fault and the Bessala fault trending S-E. The mining area belongs to the Kuludeneng syncline bounded by the Markakuri fault in the southwest and the Bessala fault in the northeast. The complex is generally distributed in NE direction, the strata of the Upper Devonian Qiya group are outcropped in the oblique core, and the wing is the Ashele group of the Middle Devonian. The mining area is in the southwestern wings of the synclinal slope. The local tectonic stress field, generated by the strike deflection of the Markakuri grave fault, resulted in a series of axial near N-S secondary fold structures (Yang et al., 2016).

Due to high magmatic activities in the Ashele basin, volcanic rocks, intrusive rocks, and dikes were developed (Fig. 1). The rock mass is distributed along with the eastern and western sides of the Markouri fault in the Qiongkur-Abagong and the Ertix fold belts. The region is well-known for non-ferrous metals, precious metals, rare metals, and muscovite deposits. The distribution of minerals is closely related to the regional stratigraphic structure, bounded by the Markakuri fault: the northeast part is the Guikan metallogenic belt and the southwest part is the Ertix metallogenic belt. Along the Qiongkur-Abagong fold belt, the Ashele copper-zinc deposit, the Karyinbura copper-zinc deposit, the Tiemoult lead-zinc deposit, the Qiaxia copper deposit, the Monku iron

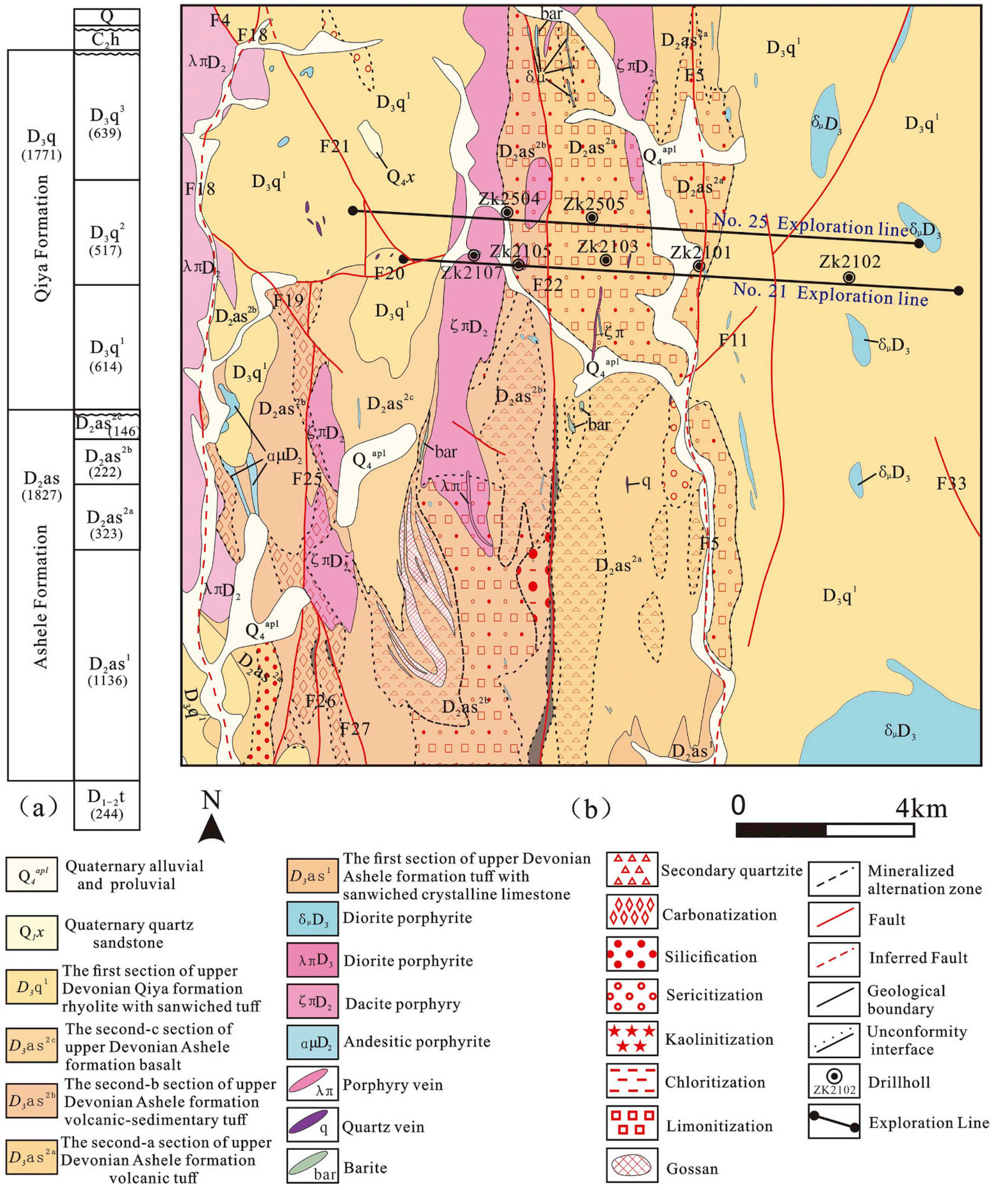


Fig. 2. (a) Stratum histogram of study area, (b) geologic map of the copper-zinc deposit.
(Revised according to the fourth Geological Battalion of Xinjiang Bureau of Geology and Mineral Resources Exploration and Development, 2002).

ore deposit, and the Kokotal lead-zinc deposit were developed, forming the Crane Hercynian metallogenic (Cu, Pb, Zn, Au, Ag, and Fe) belts.

2.2. Deposit geology

The strata of the mining area mainly comprise the Middle Devonian Ashele group and the Upper Devonian Qiya (D_3q) group (Fig. 2). The middle Devonian Ashele formation (D_2as) is geographically located in the middle, north, and southwest of the mining area (Fig. 2b). According to the lithologic differences, the Ashele formation is divided into two groups (Fig. 2a): the first lithologic group (D_2as^1) is less distributed in the southeast of the mining area, dominated by tuff, and intercalated with crystalline limestone; the second one (D_2as^2), exposed in the middle part and northwest of the mining area, is in conformable contact with the first section.

The rock assemblage for the second (D_2as^2) group is based on three layers: the intermediate-acid pyroclastic rocks lower layer (D_2as^{2a}), dominated by tuff and tufaceous breccia rocks; the intermediate-acid pyroclastic rocks and volcanic-sedimentary clastic rock middle layer (D_2as^{2b}), including minor basic lava, crystalline limestone, sericite phyllite, iron-bearing siliceous rocks, polymetallic sulfides, and barite deposits; the upper layer (D_2as^{2c}) characterized by basalt, and intercalated with a small amount of tuff and metal sulfide slice ore beds. The middle layer (D_2as^{2b}) contains a mineralized alteration area that corresponds to a wider ore-bearing zone compared to the lower layer where a number of single smaller mineralized areas form a less extended zone. In this study, we focused on the Ashele No. 1 copper-zinc deposit which is located in the middle layer (D_2as^{2b}) of the rock assemblage.

Fault structures are widely developed in the mine range, and the Markakuri fault is a first-order structural zoning fault (Fig. 1), playing an important role in controlling both the regional structures and lithofacies. The Markakuri deep fault occurred before mineralization led to the generation of a series of secondary fault structures. The Bessala fault is a typical second-order fault of the Markakuri fault. According to the direction of the secondary faults, the structural lineaments in the area can be divided into four-direction groups: N-S, N—W, N-E and E-W. The first and second groups are well-developed. The N-S trending and

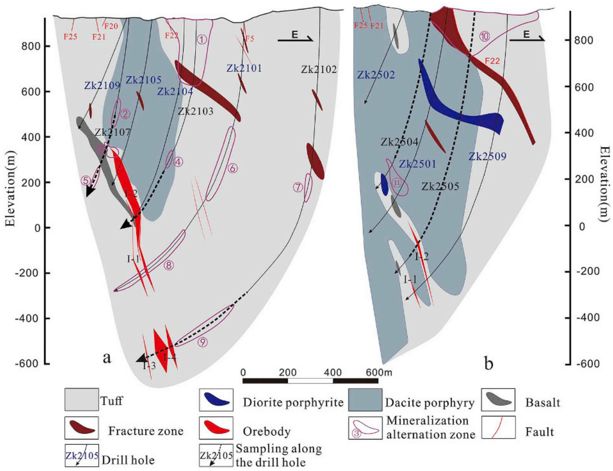


Fig. 3. Geological map of the exploration lines No. 21 (a), 25 (b).

N—W trending fault groups have a compression-torsional structure, showing the characteristics of multi-stage activity. The N-E direction and E-W direction groups are small-scaled, and they are poorly developed. Moreover, some faults are accompanied by mylonites, gneisses, flow cleavage, and other ductile deformations.

The fracture zone and mineralized alternation zone are generally developed in the fault zone, in which there are varying degrees of cataclastization, kaolinization, sericite alteration and chlorite alteration. Surface mineralization and alteration widely occur in the mining area under the influence of dynamic and hydrothermal metamorphisms. Most of the mineralized alteration zones are delineated in the mining area, which is distributed in the direction of N—W and near N-S, being consistent with the strike of strata (Fig. 2b).

The Ashele No. 1 copper-zinc deposit, distributed in S—N direction, occurs between basalt and tuff and integrates with the strata. Elevation of the orebody ranges from −600 to 888 m and the thickness varies between 5 and 120 m (Fig. 3) (Yang et al., 2016). The massive sulfide copper-zinc deposit, dominated by volcanic exhalation-sedimentary genesis and superimposed by the metamorphism and hydrothermal fluids in the later stage, shows the common characteristics of a typical submarine volcanic-exhalative-sedimentary deposits.

The ore metal minerals are mostly pyrite, chalcopyrite, and sphalerite, followed by tetrahedrite, galena, and other metal sulfide minerals. Gangue minerals are mainly quartz, sericite (muscovite), and contain a small amount of chlorite, barite, calcite, dolomite, rutile, sphene. The texture of minerals is mainly crystalline and granular, accompanied by metasomatic and metasomatic textures simultaneously. The structure of ore is mainly massive and strip structures, followed by disseminated-banded and veinlets-disseminated structures with a tiny amount of breccia structure (prospecting reports by the Ashele Copper Corporation Limited, 2016).

3. Materials and methods

3.1. Sample collection and trace-element analysis

In this work, a total of 404 rock (primary halo) samples were collected from 5 boreholes (2102, 2013, 2107, 2504 and 2505) in No. 21 and No. 25 cross-sections of the mine range (Fig. 2b). The “continuous pick-up method” was adapted for sampling collection with a distance of 5 to 10 m between the two adjacent samples. The samples were transported to the laboratory of the Southwest Metallurgical Geology Test Institute in China and analyzed for Cu, Zn, As, Sb, Bi, Sn, Ag, Pb, Mo, Ba, Co, and Au concentrations. Atomic Fluorescence Spectrometer (AFS) was utilized to detect As, Sb and Bi using nitric acid, hydrofluoric acid, and perchlorate solution. Au, Ag, Pb, Mo and Sn were measured by the Emission Spectrometry (ES). The rest of the chemical elements were determined via inductively coupled plasma atomic emission spectroscopy (ICP-AES).

The analytical detection limit is 10 ppm for Ba and Zn, 1 ppm for Co, Cu and Sn, 0.2 ppm for Mo and As, 0.04 ppm for Sb and Bi, 4 ppm for Pb, 0.035 ppm for Ag and 0.5 ppb for Au. The precisions (RSD%) of the analysis per element based on replicate measurements are 4.3% for Ba, 4.8% for Co, 4.2% for Cu, 3.88% for Zn, 4.1% for Pb, 2.5% for Mo, 4.5% for Sn, 3.8% for Ag, 4.7% for Au, 3.3% for As, 4.7% for Sb and 3.3% for Bi.

3.2. Axial zoning characteristics of the primary halo

Primary haloes of mineral deposits, which are the result of interactions between country-rocks and mineralizing fluids, are characterized by element/metal enrichment/depletion and/or mineral alterations (Carranza, 2012). The primary halo zoning of a deposit can reflect the spatial accumulation of elements during the metallogenic process, which is of great significance for prospecting and deep and concealed orebody predictions (Xing and Liu, 2011). In general, the investigation conducted for studying primary halos includes the zonality sequence and concentration of the vertical axis. Methodologies used for calculating an axial-zoning sequence of the primary halos are several including the Grigorian method and many more (Beus and Grigorian, 1977; Luo et al., 2007; Shao, 1997; Xing and Liu, 2011).

3.2.1. Zonality sequence of halo elements

The Grigorian method has been successfully applied to several mining areas in China (Chen et al., 2016; Dai et al., 2017; Li et al., 2016b; Pang et al., 2016) and, in this paper, an improved version of the same method (Yousefi and Carranza, 2015; Wang et al., 2007) was used.

Specifically, the Improved Grigorian Zoning Index (IGZI) method normalizes raw data (D) to put them into the [0, 1] interval by Eq. (1):

$$D_{ij} = \frac{\mathrm{C_{ij}} - \mathrm{Min}\, C_{ij}}{\mathrm{Max}\, \mathrm{C_{ij}} - \mathrm{Min}\, \mathrm{C_{ij}}} \tag{1}$$

where $\mathrm{C_{ij}}$ represents the concentration of j element into the i section of the elevation range.

After normalization, the data of each borehole are then divided into several (n) elevation segments ($n > 3$) and, then, the average value of the normalized data in each elevation segment is calculated (one-by-one) to determine the zonality index (E) (Table S.2, Supporting material) according to Eq. (2):

$$E_{ij} = \frac{D_{ij}}{\sum_{i=1}^{n} D_{ij}} \tag{2}$$

The highest value in different elevation segments of each element, from upper to lower, is considered as the zoning.

Since there may be several elements, which have the highest value, in the same elevation segments simultaneously, the axial-zoning sequence (Table 1) is distinguished by calculating the variation gradient (F) by means of Eq. (3):

$$\mathrm{F}_{ij} = \sum_{i=1}^{n} \frac{E_{ij}}{Max\ D_{ij}} \tag{3}$$

The smaller the variation gradient, the more it is enriched to the shallow part. The IGZI method avoids the effect of artificially unifying the original data into the same order of magnitude, and the calculations are not as complicated as the original Grigorian method.

3.2.2. Concentration zoning of primary halo

The concentration zoning is a crucial feature of primary halos. Abnormal halos in ore deposits are mainly caused by the infiltration and diffusion of the ore-bearing hydrothermal fluids, which belongs to

Table 1
Axial-zoning sequence of primary halo in selected boreholes along No. 21 and No. 25 exploration lines of Ashele copper-zinc deposit.

Exploration lines	Drilling No.	Primary halo axial zoning sequence[a] (from top to bottom)	Altitude (m)
No. 21	ZK2102	Mo-Ba-Co-Sn-(Pb—Ag)-Sb-(Au-Cu-Bi-Zn)-As	− 300 to − 600
	ZK2103	Zn-Co-Ba-(As-Bi-Sb-Cu-Sn-Mo)-(Pb-Au-Ag)	250 to 0
	ZK2107	(Au-As-Ag)-Pb-Ba-Cu-Co-(Sb—Bi)-Sn-Mo-Zn	500 to 150
No. 25	ZK2504	(Sb—Pb)-(Mo-Bi-Zn-Cu-As)-Sn-Ba-Co-(Ag—Au)	950 to 250
	ZK2505	Ba-(Bi—Sn)-(As—Sb)-(Cu—Pb)-Ag-Mo-Zn-Co-Au	950 to − 150

[a] Note: the elements in parentheses are arranged from small to large after calculating the variation gradient of the elements in the same elevation segment.

the zoning of hydrothermal precipitation, and the halo zone is the result of leachate-precipitation (Luo et al., 2007). The source of metallogenic material which forms the orebody is related to volcanism (Liu et al., 2008). By dividing the concentration of elements into zonation, the association and relative areal enrichment of elements can be distinguished as the basis for deep prediction of ore deposits (Chen et al., 2017).

In this study, for element concentration zoning, the data from the five boreholes along No. 21 and No. 25 exploration lines were used to generate the distribution maps of the axial concentration zoning of primary halo by means of a geochemistry dedicated GIS software (GeoDas) through the multifractal inverse distance weighting (MIDW) method (Cheng, 2003) (Figs. 3, 4).

To set the concentration intervals in the MIDW maps, the Concentration-Area (C-A) plot method (Cheng et al., 1994; Cheng, 1999) was applied to the interpolated grids by means of the ArcFractal plugin developed by Zuo and Wang (2019) for the ArcGIS software. The C-A fractal model has been extensively applied to separate geochemical anomalies from backgrounds and it has been feasible to produce reliable outcomes in several local and regional studies (Albanese et al., 2007, 2010, 2013, 2015; Daya Sagar et al., 2018; Lopes and Betrouni, 2009; Xie et al., 2008; Zhao et al., 2011; Zuo, 2012, Zuo, 2011, Zuo and Wang, 2016).

Specifically, for the purposes of this work, the concentration values of each grid were separated into three intervals (outer, middle and inner band, respectively) by means of strike lines at the most evident inflection points in the C-A plot (Table 2).

3.3. Statistical analysis

To analyze the elemental distribution and to unravel the relationship between metallogenic elements and other indicator elements,

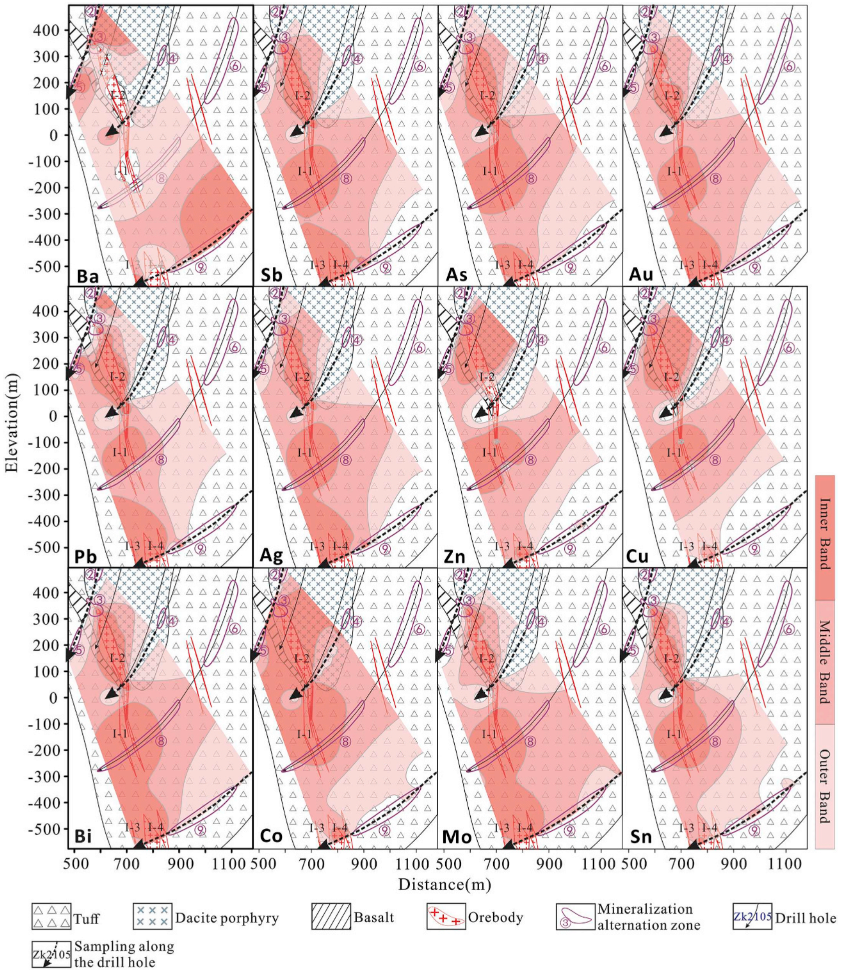


Fig. 4. Axial concentration zoning of primary halo in the cross-section of No. 21 exploration line.

Table 2
Geochemical cutoff values of rock samples (n = 404) classified by C-A fratal model.

Cross section	Exploration line No. 21			Exploration line No. 25		
Elements	Outer band	Middle band	Inner band	Outer band	Middle band	Inner band
As	32.77	151.46	240.92	2.14	2.81	4.53
Sb	3.13	6.44	11.49	0.19	0.26	0.32
Mo	3.00	5.35	8.81	0.42	0.57	0.91
Zn	575.22	929.79	1459.30	61.70	78.06	88.12
Bi	1.08	4.88	8.43	0.07	0.14	0.25
Cu	1915.23	3453.23	7190.09	25.88	36.45	60.39
Au	10.20	37.15	75.29	0.47	0.71	1.27
Sn	3.61	4.46	5.95	0.51	0.70	0.80
Ag	0.25	0.45	0.71	0.05	0.06	0.09
Pb	24.87	37.97	44.45	5.51	6.68	7.49
Co	7.82	12.92	16.68	6.33	15.14	16.55
Ba	158.78	248.57	314.12	99.22	133.67	176.78

Note: Au values are multiplied by 10^{-9}, concentrations of the other elements are multiplied by 10^{-6}.

univariate (Table 3) and multivariate statistical analyses were performed using geochemical data from a total of 404 rock samples. To apply the concept of dimension reduction to categorize variables into a few comprehensive factors regarding the internal dependence of the correlation matrix of the original variables (Field, 2013; Reimann, 2008), an R-mode factor analysis (FA) was completed after applying a log-transformation to geochemical data for normalization purposes (Albanese et al., 2013). The fitness of the elements was tested for FA and Ba and Co were excluded from the analysis since their low correlation with the other elements in the dataset (always < 0.6) and Sb was eliminated because of its low communality extraction value (< 0.7), as well.

After that, the calculated Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) (0.901) revealed that the data met the requirements for FA. After a varimax rotation with Kaiser Normalization, two components were extracted by means of the principal component analysis method (Table 4).

3.4. Calculation of geochemical prediction indication

Geochemical prediction indication (Vertical Element Contrast Coefficient - VECC, Geochemical Parameter - GP and trilinear diagram) is a useful approach for discriminate against the negligible possibility for evaluating the potential of an orebody at depth (Li et al., 2016c). Both the VECC and GP intuitively demonstrate the fluctuation (spatial variation) of signal element or elements association on the vertical profile, while the trilinear diagram better illustrates the association between orebody and mineralized alternation zone in different elevation segments.

In this paper, data (Table S.2, Supporting material) were selected to form the VECC plot, and the value of GP and trilinear diagram were calculated through Eq. (1).

The vertical GP, closely reflecting the development degree of the supra-ore halo relative to the sub-ore halo, was calculated by the following Eq. (4):

$$GP = \frac{w(As + Sb)}{\mathrm{w(Bi + Mo + Sn)}} \tag{4}$$

where the value of elements inside the parentheses was calculated by means of Eq. (1).

The larger the parameter value (being characteristics of the supra-ore halo of the orebody), the greater is the possibility of the concealed orebody in the depth (Zhang et al., 2010).

To better interpret the hidden characteristics of primary halo, a trilinear diagram of primary halo was also plotted for each borehole (Fig. 10). The normalized dataset of supra-ore halo elements (As and Sb), near-ore halo elements (Cu, Pb and Zn) and sub-ore halo elements (Bi, Mo and Sn) was combined to three groups, respectively, and the value of each axis of the diagram was defined by the arithmetic sum of each group. The areal distribution characteristic of primary halo can be recognized by delineating (Value > 40%) favorable supra-ore (G-area), near-ore (Y-area), and sub-ore halo (R-area) spatial domain on the diagram.

All geochemical prediction indications aforementioned can be applied as a forceful criterion to interpret the overlap of the primary halo model (Figs. 11 and 12) and primary halo prospecting in the depth.

Table 3
Statistical parameters of rock samples (n = 404) from the Ashele mining reigon.

Element	Unit	Min	Max	Mean		SD	Skewness	Kurtosis
				Statistic	SE			
As	mg kg^{-1}	0.59	2620.70	30.87	8.97	180.22	11.30	141.88
Sb	mg kg^{-1}	0.03	130.46	1.61	0.45	9.00	10.38	125.02
Bi	mg kg^{-1}	0.01	79.99	1.19	0.30	6.12	9.61	103.73
Sn	mg kg^{-1}	0.35	22.27	1.46	0.11	2.17	4.81	32.96
Ag	mg kg^{-1}	0.02	5.00	0.17	0.02	0.46	7.05	61.52
Pb	mg kg^{-1}	2.21	500.00	14.48	1.78	35.79	8.54	95.22
Mo	mg kg^{-1}	0.07	46.62	1.92	0.22	4.47	4.99	33.40
Ba	mg kg^{-1}	9.64	1849.90	224.52	11.33	227.71	2.91	11.88
Co	mg kg^{-1}	0.79	69.72	15.47	0.42	8.38	1.26	5.02
Cu	mg kg^{-1}	1.58	29,740.00	244.16	80.94	1626.89	15.47	271.69
Zn	mg kg^{-1}	21.31	19,100.00	161.19	47.70	958.83	19.25	380.40
Au	ug kg^{-1}	0.25	1440.00	10.02	4.21	84.56	14.49	225.97

Table 4
Varimax-rotated factor analysis of the selected variables for rock samples of the exploration lines.

Variables	Factor		Communalities
	F1	F2	
As	**0.647**	0.594	0.772
Bi	**0.899**	0.239	0.865
Sn	**0.903**	0.135	0.834
Ag	**0.657**	**0.614**	0.809
Pb	0.397	**0.825**	0.838
Mo	**0.793**	0.355	0.755
Cu	**0.740**	0.399	0.707
Zn	0.077	**0.867**	0.757
Au	0.584	**0.689**	0.816
Eigenvalues	68.227	11.248	
Total variance in %	46.174	33.301	
Cum. of total variance (%)	46.174	79.475	

Bold entries: loading values over 0.60.

4. Result and discussion

4.1. Primary halo axial-zoning sequence

The main reasons for the primary halo zoning during the magma migration are a gradual evolution of the metallogenic fluid properties, and the various conditions generating a range of minerals in the form of different assemblages (Shao, 1997). After studying a large number of volcanic hydrothermal copper-lead-zinc polymetallic deposits, Liu and Ma (2007) and Liu et al. (2008) suggested that, for this kind of deposits, the supra-ore halo elements are I, Hg, As, Sb, B and Ba; the near-ore halo elements are Pb, Zn, Ag, Au, Cd, Mn, and Cu; the sub-ore halo elements are Cu, Bi, Co, Mo, W, and Sn.

In the present study, after processing the data of 12 elements, we determined that the supra-ore halo elements in Ashele deposit are As, Sb, Ba, and Au; the near-ore halo element assemblage is Pb, Zn, Cu, and Ag (Au), and sub-ore halo element assemblage is Bi, Co, Mo, and Sn. This can be a criterion for judging whether there is a reverse element zoning in the calculated zoning index, as well.

The final axial-zoning sequence of the primary halo was, also, determined as reported in Table 1. Through the investigation of the areal distribution of different temperature elements, associated with the different halos, this method can result in an effective tool to predict and evaluate the occurrence of the deep orebodies.

Both mineralized alteration zones and orebodies can form primary halo anomaly. Based on this theory and combined with information proceeding from the geological profiles (Fig. 2), the calculated axial-zoning sequence of each borehole (Table 1) were divided into different groups (A, B, C). Aiming at a better understanding of the axial-zoning sequence, two cross-sections were produced to explore the axial-zoning characteristics of primary halo (Table 5). All boreholes indicate the existence, at least, of three groups of primary halo zoning. Although a quantity of the primary halo anomalies can be related to orebodies and/or mineralized alteration zones, some of them remain still unexplained. According to Table 5, there is a superposition of sub-ore halo and supra-ore halo at the end of boreholes 2102, 2103 and 2505, revealing a series of small known orebodies beneath the main orebodies. Besides, some anomalies in borehole 2504 and 2107 or even Zk2505 do not correspond to the orebody or alteration zone, indicating a high possibility of concealed blind ore bodies.

4.2. Primary halo concentration zoning

Regarding the axial concentration zoning of primary halo in No. 21 exploration line (Fig. 4), most of the elements (except for Ba and Co) develop a clear inner anomaly zone which corresponds well to the ore deposits, while the outer and middle zone anomalies are not clearly explainable. Among all the elements, Cu, Zn (Pb) are the main metallogenic and near-ore halo ones, whose anomalies in sections are most related to orebodies. Represented as associated beneficial components, both Au and Ag are supra-ore halo elements, with axial concentration zoning anomalies similar to the ones of the above near-halo elements. Since Sb and As are supra-ore halo elements, they have good abnormal continuity in the whole axial zoning, and the anomaly scale is wider than that of Cu and Zn. It can better reflect the characteristics of strong diffusivity and high mobility of low-temperature elements (As and Sb). Bismuth, Mo, and Sn have similar vertical distribution showing large-scale sub-ore halo anomalies at the elevation of −300 to 0 m, possibly caused by the orebodies or No. 8 mineralized alteration zone, and the relatively undeveloped anomalies of orebodies at the elevation of 0 to 200 m. Moreover, Co (as a high-temperature element) is representative of a large inner anomaly zone surrounding the orebody; while this anomaly zone can hardly refer to the precise location of the orebodies. The anomalous zone of Ba (low-temperature element) is small and the spatial matching with the orebody seems to be not so good. The local Ba anomaly, probably, resulted from the barite alternation zone, indicating the presence of both a mineralized alteration zone and a fracture zone as reported by the Ashele Copper Corporation Limited (2016).

The axial concentration zoning of the primary halo in the No. 25 (Fig. 5) is weaker than the No. 21 exploration line (Fig. 4). However, the definition of the anomaly and the axial zoning of elements are higher and more distinguished in the cross-section drawn for the No. 25 exploration line. The fracture zone, which is rooted from the surface borehole ZK2504, and which also extended from the east side towards the deep underground, was intercepted by borehole Zk2505 at an elevation of 700 m (Fig. 2b). The fracture zone was responsible for a significant anomaly of most of the elements in the elevation of 600 m and 900 m. Cu, Zn and Pb, the medium-temperature near-ore halo elements, are affected by the fracture zone and developed distinct large-scale inner zone anomalies at the elevation of 600 m and 900 m. Furthermore, small-scale inner Pb anomalies occur between −100 and 50 m, and the Cu anomalies are limited developed in-depth. Zn and Cu, as the major metallogenic elements, along with Mo and Ag, are characterized by a series of inner anomalies between 300 and 500 m; these anomalies may have been caused by the presence of intrusive rocks. As near-ore halo elements, Au and Ag have similar anomaly patterns, in which, at the level of −100 m to 50 m, Au develops distinct anomalies in correspondence with the I-2 orebody and Ag a small outer anomaly. From the vertical distribution of primary halo elements in hydrothermal deposits, it can be inferred that the precipitation temperature of Ag is higher than that of Au, which indicates that the latter has stronger mobility than the first one (Shao, 1997). Arsenic, Sb and Ba are supra-ore halo elements showing remarkable anomalies at the elevation of 700 to 900 m, close to the surface, and they may have been subjected to near-surface weathering and denudation. Moreover, Ba, Sb and As have good inner zone anomalies at the level of −100 to 50 m, with a good spatial correspondence with the deeper I-1 and I-2 orebodies. Bi, Mo, and Sn, which are indicators of high-temperature sub-ore halo element, were identified at 600 to 900 m as well. The anomalies of these three elements did not extend in the depth. Similar to section of the No. 21 exploration line, the variety of anomalies, in spatial distribution, among three elements (Bi, Sn and Mo) may be caused by multistage of hydrothermal fluids which have a close relationship with fracture and mineralized alteration zone.

4.3. Factor analysis

The results of orthogonal rotation of factor analyzed were listed in Table 4. The cumulative percent value of the total variance explained was 79.47% and the following factors, including all the elements with a loading ≥0.6, were defined:

Table 5

Axial zoning characteristics of primary halo in exploration lines No. 21 and No. 25.

Borehole	A	B	C	Matching	Altitude (m)	Remarks
ZK2102	Mo			MAZ	−300 to −290	Silicification-carbonation-pyrite mineralization & chlorination
		Ba		I-4	−325 to −300	Caused by ore deposit
	Co-Sn			MAZ	−475 to −375	Silicification-carbonation-pyrite mineralization & chlorination
		Pb-Ag		I-4	−500 to −475	Caused by ore deposit
		Sb		I-4	−550 to −500	Caused by ore deposit
		Au-Cu-Zn		I-4	−550 to −525	Orebody exposed section
		Bi		I-4	−550 to −525	
			As	I-3	−576 to −550	
ZK2103	Zn			MAZ	225 to 250	Chlorite alteration zone
	Co			MAZ	200 to 225	
		Ba		I-2	75 to 100	Caused by ore deposit
		As		I-2	50 to 75	Orebody exposed section
		Bi		I-2	50 to 75	
			Sb	I-1	50 to 75	Indicates the small orebody I-1
		Cu		I-2	50 to 75	Orebody exposed section
		Sn-Mo		I-2	50 to 75	
			Pb-Ag-Au	I-1	25 to 50	Indicates the small orebody I-1
ZK2107	As			I-2/MAZ	475 to 496	Indicates the superposition area of I-1 and alteration zone
	Au-Ag-Pb			I-2/MAZ	425 to 496	
	Ba			I-2/MAZ	400 to 425	
	Cu			I-2/MAZ	350 to 375	
		Co		MAZ	325 to 350	There is a sub-ore halo caused by sericite and chlorite alteration, and basalt is exposed in the middle 275 to 350 m, or the near-ore halo and sub-ore halo are not developed, and the tail may indicate an unknown orebody/mineralization zone
			Sb	UNKNOW	250 to 275	
		Bi-Mo-Sn		MAZ	175 to 275	
			Zn	UNKNOW	152 to 175	
ZK2504	Sb			DP	700 to 750	Drilling through diorite porphyrite may intersect the halo formed by diorite to form a symmetrical halo structure.
	Pb			DP	700 to 750	
	Mo-Bi			DP	650 to 700	
	Zn-Cu			DP	650 to 700	
	As			DP	650 to 700	
		Sn		UNKNOW	500 to 550	The area is a prospecting area, indicating an unknown orebody/mineralization zone
		Ba		UNKNOW	400 to 450	
		Co		UNKNOW	300 to 350	
			Ag-Au	MAZ/DP	288 to 300	A small amount of pyrite mineralization
ZK2505	Ba			MAZ&FZ	900 to 920	Located in the tectonic fracture zone, where pyrite mineralization and silicification alteration are developed.
	Bi-Sn			MAZ&FZ	850 to 900	
	As-Sb			MAZ& FZ	750 to 800	
	Cu-Pb			MAZ& FZ	650 to 700	
		Ag		UNKNOW	400 to 450	Prospecting area (lack of near-ore halo)
		Mo		UNKNOW	350 to 400	
		Zn		UNKNOW	300 to 350	
		Co		I-2	−50 to 0	Indicates orebody I-2
			Au	I-1	−100 to −50	Indicates orebody I-1

Note: MAZ = mineralized alteration zone, FZ = fracture zone, DP = diorite porphyrite; black = supra-ore halo element, green = near-ore halo, red = sub-ore halo in Group A, B, C.

- F1: Sn, Bi, Mo, Cu, Ag and As
- F2: Zn, Pb, Au and Ag

Both factors have a good association with the mineralization process while F2 element assemblage count for a great proportion of possibility. F1 shows high loadings for high-temperature indicator elements (Mo, Sn and Bi) and metallogenic elements (Cu and Ag), and low-temperature indicator elements (As). F2 includes four metallogenic elements (Zn, Pb, Au and Ag). Previous studies (Chen, 1996; Gao et al., 2010; Wang, 1995; Wu et al., 2015; Ye et al., 1998) showed that after the formation of layered main orebodies during the exhalative-sedimentation period, the Ashele copper-zinc deposit undergone multistage of reformations (such as metamorphism, magmatic-hydrothermal superposition, and epigenetic oxidation). Ye et al. (1997, 1998) found that the metallogenic temperature of the hydrothermal sedimentary stage of the Ashele deposit is around 80 to 310 °C, the formation temperature of the massive sulfide layer ranges from 80 to 250 °C, and the formation temperature of stockwork and disseminated deposit is between 230 and

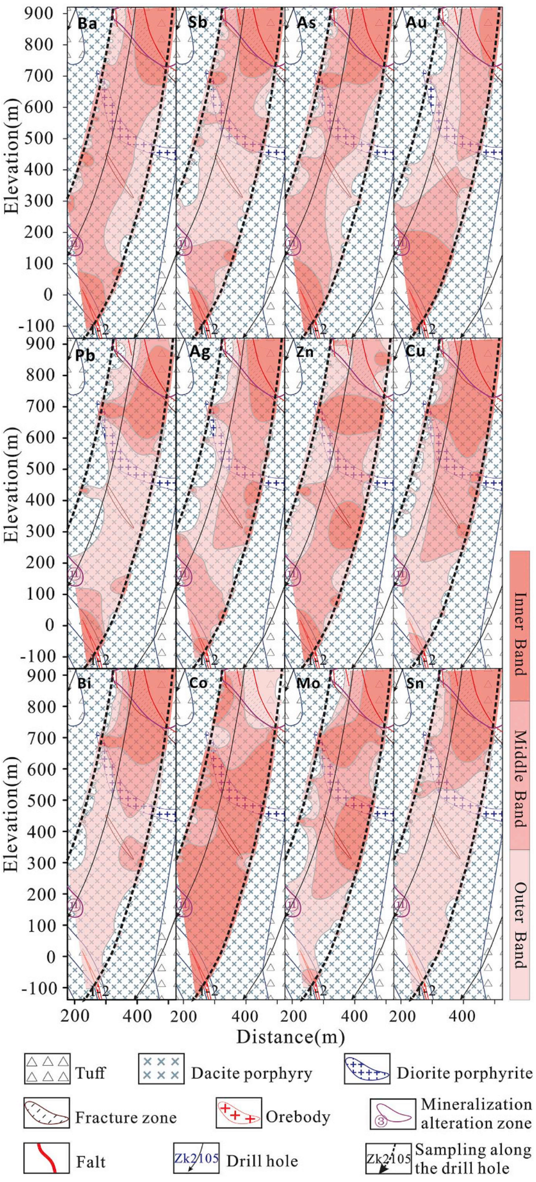


Fig. 5. Axial concentration zoning of primary halo in the section of No. 25 exploration line.

310 °C. Chen Yuchuan. (1996) demonstrated that during the period of metamorphic reformation and hydrothermal superposition, the hydrothermal temperature increased and varied from 310 °C and 540 °C. Concerning the difference of the extracted principal factors (Table 4) and the temperature of ore-forming fluids in the mine range, it can be concluded that the element association in F1 is mostly related to metamorphism and hydrothermal superposition. On the other hand, the element assemblage of F2 mainly explains the mineralization of exhalative-sedimentation deposits. Furthermore, both two factors containing metallogenic elements (e.g., F1: Cu (Ag); F2: Zn (Pb)), more or less, show the multistage- superposition genesis of deposit during the ore-forming process.

4.4. Vertical variation of the factor scores

Mapping the factor scores is frequently applied in exploration geochemistry to represent the spatial variation of different geological processes. Higher absolute value of the factor score indicates a greater impact of the geological or geochemical processes (Wei et al., 2011).

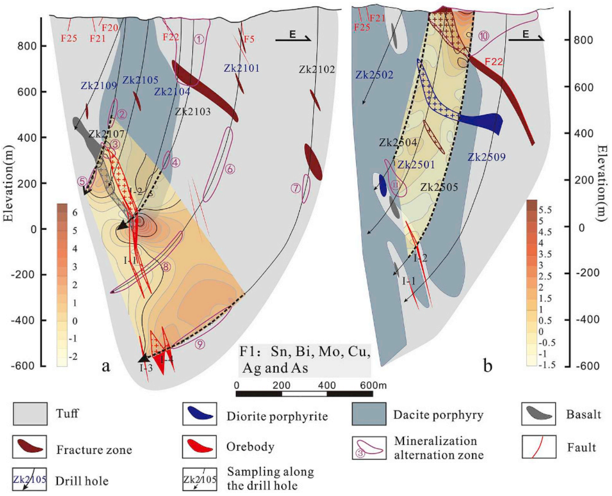


Fig. 6. Interpolated score maps for Factor 1 (F1): a) cross-section along the No. 21 exploration line; b) cross-section along the No. 25 exploration line. The plus signs (+) in diorite porphyrite, basalt, fracture zone and orebody represent the overlapping areas with the interpolation map of these geological bodies.

4.4.1. F1 (Sn, Bi, Mo, Cu, Ag, As)

Vertical variation of F1 map (Fig. 6) reveals that most of the high scores correspond to the mineralized alteration zones. According to Fig. 6a, the elevated factor score values between −500 and −300 m correspond to the No. 9 mineralized alteration zone. Same-type alteration (Carbonate minerals, pyrite mineralization and silicification) are observed both in No. 8 and No. 9 mineralization zones (prospecting reports by the Ashele Copper Corporation Limited, 2016), indicating No. 9 mineralization zone might be extended up to No. 8 mineralization zone, which may be the (one) reason of considerable association between the elements of F1 appears on the lower right of the orebody (I-2) at sea level (Fig. 6a). Nevertheless, there is a lack of chemical data for the No. 8 mineralization zone and further investigation is needed to confirm this assumption.

High scores (Fig. 6b) detected in the No. 25 exploration section coincide with the No. 10 mineralization fracture alteration zone. The Sn-Bi-Mo assemblage in F1 can be attributed to the high-temperature silicification, one of the most common alterations in the wall rocks (prospecting reports by the Ashele Copper Corporation Limited, 2016). Further, the pyritization or micro-aggregates of Ag and As sulfides in the country-rock can explain the presence of As in F1. The copper-bearing minerals (e.g., chalcopyrite and tetrahedrite) are mainly xenomorphic granular and commonly occupied in the margins or fissures of pyrite in the orebody, while they are mostly in gangue minerals in the alteration zone (Chen, 1996). Hence, copper enrichment, as the main metallogenic element, both in the sulfide orebodies and the wall rock alteration zone can be well explained. This is also the reason why Cu and As have similar, wide, anomaly pattern in vertical concentration distribution (Figs. 4 and 5), due to the same interdependent relationship with pyrite. On the other hand, silver-bearing minerals are mostly distributed in copper-zinc polymetallic sulfides, resulting in an approximate vertical primary halo distribution of Cu and Ag. It can be seen that high loading values of Sn, Bi and Mo characterize the alteration zone better than Cu, As and Ag which are the indicators for the orebody.

4.4.2. F2 (Zn, Pb, Au, Ag)

F2 map (Fig. 7) depicts distribution patterns of scores clearly influenced by the presence of ore bodies. As a matter of the fact, in the cross-section of the No. 21 exploration line (Fig. 7a), elevated factor scores between 400 and 600 m b.s.l. overlap the known orebody (I-3, I-4); In addition, moderately high scores are uniformly distributed on

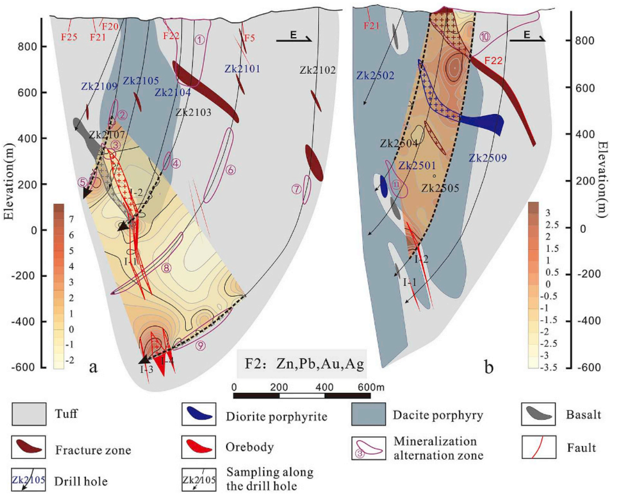


Fig. 7. Interpolated score maps of Factor 2 (F2): a) cross-section along the No. 21 exploration line; b) cross-section along the No. 25 exploration line. The plus signs (+) in diorite porphyrite, basalt, fracture zone and orebody represent the overlapping areas with the interpolation map of these geological bodies.

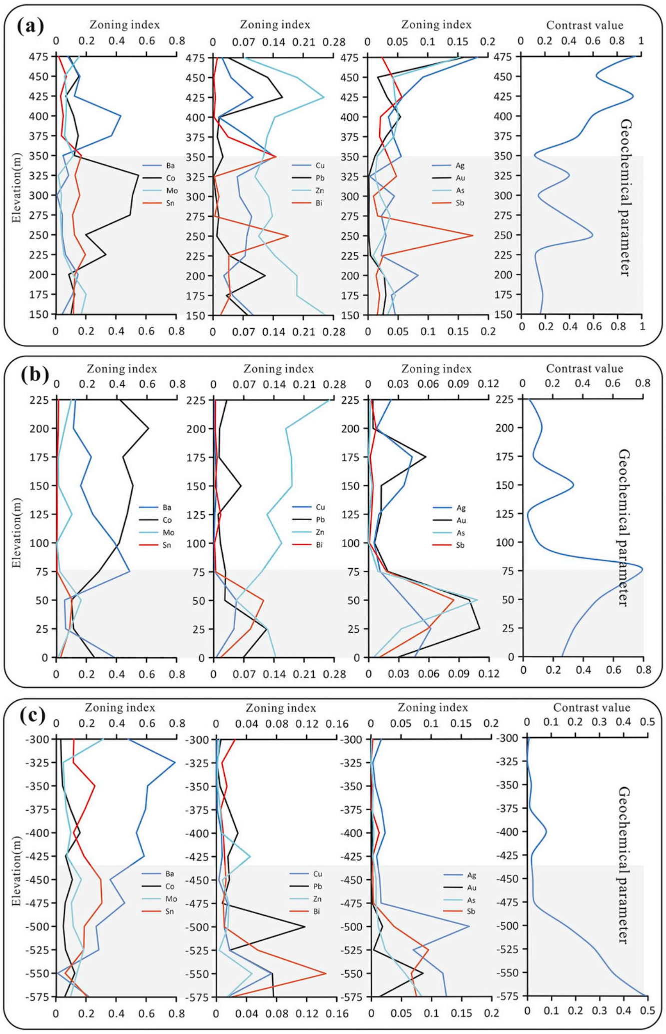


Fig. 8. Zoning index of primary geochemical anomalies of Zk2107 (a), Zk2103 (b), Zk2102 (c) along No. 21 exploration line. The shadow area is the spatial distribution of orebody.

both sides of the deposit (I-2) at 0 m elevation and low values correspond well to the mineralized alteration zone. Similarly, the strong association of Zn, Pb, Au, Ag at the elevation between −150 and 200 m (Fig. 7b) is related to the spatial distribution of the orebody (I-2) along the No. 25 exploration line. Furthermore, the high scores between boreholes ZK2501 and ZK2505 between 600 and 700 m a.s.l. suggest that the area is a potential metallogenic source.

Zn, with the highest loading in F2, is another main ore-forming element in the Copper-Zinc mining district, and its ore mineral (sphalerite) is mainly hosted in copper-sulfur orebodies, while galena is closely symbiotic or embedded with sphalerite. Additionally, along No. 21 exploration line, where the orebodies (I-1, I-2, I-3 and I-4) are located, Cu and Zn (which are the main metallogenic elements) show a similar spatial distribution pattern for their anomalies in contrast with Pb which is characterized by an extension of the anomalous inner band smaller than Zn (Fig. 4). On the other hand, in No. 25 (Fig. 5) exploration line, in which mineralized alternations are widely distributed, Zn features anomaly patterns other than Cu and Pb, (especially in the top-most part of the section) suggesting that both Cu and Zn are elements highly related to the main orebody, while Cu can be linked to the presence of mineralized alternation zone, as well. This can also explain why two ore-forming elements (Cu and Zn) are in different factors.

The reason for an insignificant loading of Au (Ag) in this F2 probably caused by the widely distributed of them in sulfide orebodies and gangue minerals such as pyrite in the form of intermetallic between natural gold and silver. Furthermore, Ag and Pb occur in Au and Zn ore minerals as companions, respectively (Chen, 1996), and their abnormal morphological scale is poor corresponding to I-1 and I-2 orebodies (Figs. 4 and 5), which may indirectly lead to the result of no marked high scores shown in the I-2 orebody. In conclusion, most of the F2 element assemblage is metallogenic being strictly related to the orebody.

Furthermore, it is also reported by the Ashele Copper Corporation Limited (2016), in field surveys of the mining area, the formation of the metallic minerals (e.g. pyrite, chalcopyrite, tetrahedrite and sphalerite) can be characterized into two generations (the exhalative-sedimentation syngenetic generation, and the late hydrothermal deformation and metamorphic superposition generation), in which these former metallic minerals are mainly distributed in copper-sulfur orebody and the latter is mostly in fissures of gangue minerals. This is also a reason why most of the elevated scores in the F1 and F2 factor scores are distributed in the mineralized alteration zone and the orebodies, respectively (Figs. 6 and 7).

It is evident that the factor score can separate multistage ore-forming processes. Variations of the F1 factor scores towards depth (Fig. 6) are consistent with the pattern of Sn, Mo, Bi, Cu, Ag, As association in primary halo (Figs. 4 and 5), revealing that it coincides with the mineralized alteration zone formed by metamorphic transformation and hydrothermal superposition activity. On the other hand, the F2 element association Zn, Pb, Au, Ag almost indicates the formation of ore bodies by the mineralization of exhalative-sedimentation deposition.

4.5. *Vertical variation of primary halo characteristic*

A similar tendency of the VECC curve, on vertical profile, was taken into consideration to classify the 12 elements into three groups as follows: 1) Au, Ag, As, and Sb; 2) Cu, Pb, Zn, and Bi; and 3) Ba, Co, Mo, and Sn.

4.5.1. *No. 21 exploration line*

Fig. 8 depicts the variable trendlines of the VECC versus height. The ZK2107 is located on the upper left side of the I-1, I-2 orebody (Fig. 2).

the shadow area demonstrates the projection area of the orebody in vertical elevation. Peak values of the VECC revealed in the upper part (350 to 475 m) of the borehole (Zk2107) are enriched with supra-ore halo elements (Au, As, Ba) and few near-ore halo elements (Cu, Pb, Zn, Ag) (Fig. 8a). Although elevated concentrations of sub-ore halo elements (Co, Sn, Bi) are observed in the middle part (225 to 350 m), the lower part (150 to 225 m) is enriched with most near-ore halo elements (Cu, Pb, Zn). The descending trend of GP with increasing depth can be divided into three segments according to the peak trend at the elevation of 350 to 475 m, 225 to 350 m, and 150 to 225 m. Combined with the profile (Fig. 2), the first elevation segment (350 to 475 m) is above the orebody slope, in which the supra-ore halo is well developed, and a maximum value is found on the curve (Fig. 8a). Then a double-peak abnormality displays between 220 and 350 m, caused by the right-side I-2 orebody. The third elevation segment (150 to 225 m), with a trough-shape curve, corresponds to the No. 5 alternation zone, indicating a relative enrichment of sub-ore halo. The borehole Zk2103 located on the upper right side of the I-1 orebody (the shadow area of the curve is the area where the borehole it penetrates) provides valuable information on the geochemistry of the orebody as well (Figs. 2a and 8b). The VECC of supra-ore elements (As, Sb, Au and Ag) fluctuate inconsiderably and show an extremely increase at 0 to 75 m above sea level which directly corresponds to the I-1and I-2 orebodies. The sub-ore and near-ore halos elements (Zn and Co) curve show a gradually decreasing trend with increasing depth, while Cu, Pb, Bi, Mo and Sn have a comparable areal variation with As and Sb. Moreover, the GP increases gradually from shallower to deeper parts with two anomaly peaks at 150 m and 75 m, revealing copper-sulfur industrial I-2 and I-1orebody, respectively. ZK2102 has drilled below the position of the I-1, I-2 orebody, and the tail of the borehole intersects the areal distribution of the I-3, I-4 orebody (the shadow area of Fig. 8c). The VECC of supra-ore and near-ore halo elements do not vary significantly along the borehole, while several anomalies are shown at the elevation of −575 to −475 m. The VECC of Ba and Sn fall as the depth increases, and a smoothly running of VECC on Mo and Co. The GP (Fig. 8c) increase sustainably from −475 to −575 m of the profile and reaches the highest value at depth. According to the drilling profile map (Fig. 2a), high values of supra-ore halo with low values of sub-ore halo elements at depth probably suggest the I-3 and I-4 orebodies in the lower part of the I-1, I-2 orebody. The increasing rate of GP towards the end of the borehole is also another piece of supporting evidence. Besides, the insignificant peak at −425 to -375 m refers to the No. 9 mineralization alternation zone.

4.5.2. No. 25 exploration line

Two boreholes (Zk2504 and Zk2505) in No. 25 exploration line (Fig. 2), with an approximately superimposed vertical sampling elevation (250 to 900 m), where the vertical characteristic (VECC and GP) of two boreholes can be comparable, are adjacent to each other. A small amplitude fluctuation of VECC for most elements was presented along with the vertical altitude (250 to 900 m) on both boreholes. Meanwhile, a peak value of element assemblage except for Sb, Ba and Co, which correspond to the high values concentrated in Fig. 9b as well, was observed at the altitude of 650 m right beneath the No. 10 mineralized alternation zone (Fracture zone). The same result was found in factor scores (Fig. 7b), indicating a favorable prospecting area may have been here. As the depth decreases (250 to 900 m), the GP of Zk2504 and Zk2505 present as a reciprocating wave variation with comparable features, which can be explained by the multistage metamorphic-deformation and hydrothermal superimposed activities (Li et al., 2016b). This interpretation is consistent with the geological characteristic supported by the intrusive body (diorite porphyrite), fracture zone and widely-distributed mineralized alternation zone in the mining area, and the various deformation degree of rocks (prospecting reports by the Ashele Copper Corporation Limited, 2016). As for the tail part (−150 to 250 m) of Zk2505, the repaid increasing of GP which is three times higher than Zk2504 corresponds well to the I-3 and I-4 orebodies, while we cannot exclude the possibility of concealed deposit in the depth.

Both the VECC and GP of two cross-sections well reveal the spatial distribution of the orebody, in which the enrichment of the supra-ore halo at the tail of ZK2103, ZK2102 and ZK2505 directly indicate the deep copper-sulfur industrial orebody. While the ZK2107 and ZK2504 (ZK2102 and ZK2505) also indicate deep concealed orebodies or mineralized alteration zones to a certain extent. In addition, both orebody and mineralized alteration zone can cause a high VECC of supra-ore halo and an increasing tendency (or a reciprocating wave) on GP; However, the intensity of VECC and GP of the orebody is stronger than that of mineralized alteration.

4.6. Trilinear diagram of the primary halo

To better investigate the relationship between the different ore halos and the areal distribution of orebodies or mineralization zone, trilinear diagrams were generated (Fig. 10). The development of sub-ore halo in ZK2102 (Fig. 10a) is relatively high, and all the data (Point 4, 5, 6, 7, 8, 9, 10; from −500 to −350 m) being in good agreement with the No. 9 mineralized alteration zone, fall into R-area. Regarding the profile section (Fig. 2a), the elevation segments between −600 and −520 m (points 11 and 12) are in the location of I-3, I-4 orebody; According to Fig. 10b, most of the points in ZK2103 are plotted on the Y-area, indicating the near-ore halo is higher developed than supra and sub-ore halo. Only a few points (5, 8 and 9) are located in the R-area, including segments 8 and 9 (elevation of 31 to 75 m) that coincidence with the spatial distribution of I-1, I-2 orebody (Fig. 2a); Besides, most of the points in ZK2107 (Fig. 10c) fall in the Y-area and few in R-area. The segments 6, 7 and 8 (between 300 and 350 m), in which the No. 3 mineralized alteration developed, are located on the left side of I-2 orebody; and points 10, 12, 13 and 14 (elevation of 150 to 250 m) correspond to the No. 5 sericite alteration; Majorities of the points in ZK2504 (Fig. 10d) are plotted beside G-area, and few are located in the Y and R-areas, indicating that the supra-ore halo is relatively more developed. At the beginning of Zk2504 (between 850 and 912 m, Fig. 2b) where the No. 10 mineralized alteration zone is locating, points here are plotted in G-area and Mixed-area (point 2 and 1, respectively); Many samples collected from the ZK2505 borehole are classified in the near-ore halo category (Y-area, Fig. 10e). The G-area in Zk2505 includes segments (3, 4 and 17) which is an indicator for fracture zone or orebody. The points 21and 22, with a close relation to the I-2 orebody at the elevation of −50 to −139 m, fall into the G-area, indicating the occurrence of the I-1 beneath I-2 orebody. Furthermore, the surficial segments 1 and 2, in which the No. 10 mineralized alteration zone results from the fracture zone, have a well-developed sub-ore halo (R-area) with no obvious supra-ore and near-ore halo which may be destroyed by multistage tectonic activities.

Clearly, all of the mineralized alteration zones, and orebodies in No. 21 exploration line are well distributed in R-area of each borehole (Fig. 2b). It is no doubt for those falling into the R-area is an especially powerful indicator for prospecting. The segment at the end of boreholes falls into the overlapping area (superposition of supra-ore and sub-ore halo), suggesting the high possibility of the existence of an orebody or mineralized alteration zone at depth. It is consistent with the judgment criterion of the superimposed halo of the blind orebody structure proposed by Li et al. (Li et al., 1995; Li, 1998; Li and Zhang, 1999; Li et al., 2016a).

4.7. Primary halo superposition model and deep target delineation

Single mineralization forms different beaded and imbricate orebodies. When the two orebodies are close to each other, their primary halos also form a partial superposition in space. The beaded orebodies, for example, have overall supra-ore and sub-ore halos. Although sub-ore halo of the upper orebody and supra-ore halo of the lower orebody

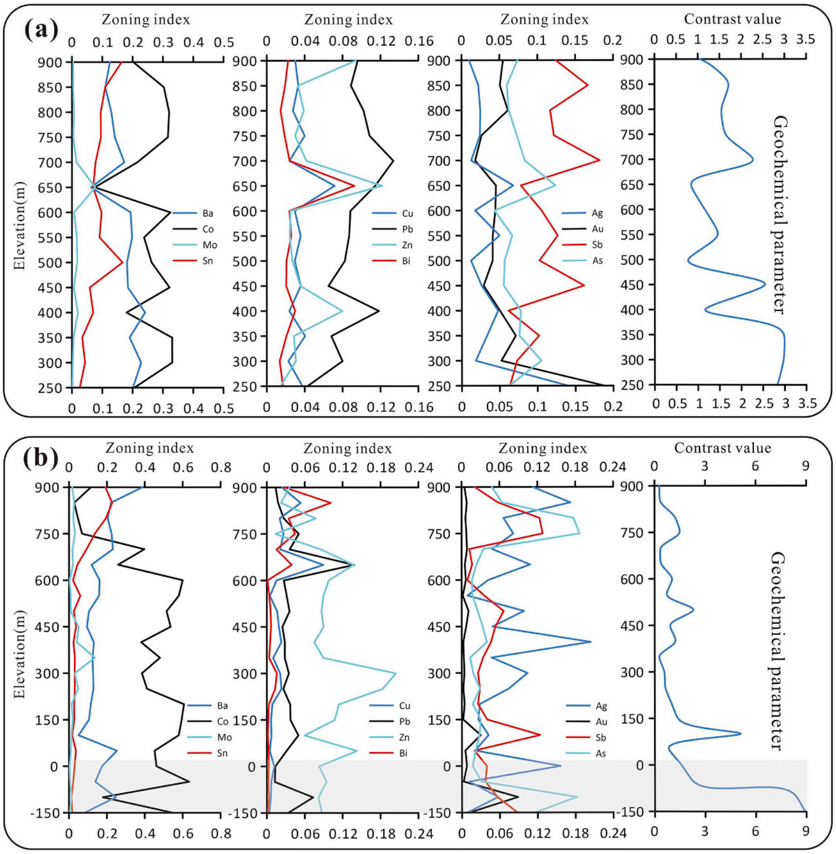


Fig. 9. Zoning index of primary geochemical anomalies of Zk2504 (a), Zk2505 (b) in No. 25 cross-section. The shadow area is the spatial distribution of orebody.

are small, the supra-ore and sub-ore halos coexist while the orebodies are close in space during the mineralization period (Li et al., 2014). On the basis of geological setting (Figs. 2 and 3), the vertical geochemical characteristics of the primary halo (Figs. 4, 5, 6, 7, 8 and 9 or Table 5) and the triangular diagram (Fig. 10), the ideal model for the superposition of the primary halo is constructed in Figs. 11 and 12. The identification standards used for developing the model are as follows:

(1) The Ashele copper-zinc deposit formed either during or at the end of the intermittent volcanic eruption (or the end of the cycle). Majority of the orebodies host near the upper and lower interlayer interface of the volcanic layer. As a result, the volcanic-sedimentary clastic rocks and interlayer interface of volcanic-sedimentary facies are leading prospecting indicators. Further, according to the characteristics of the similar deposits at home and abroad, the association of the mineralization with syngenetic alteration of wall rocks is another important prospecting indicator. Furthermore, the multistage structures, revealing multi-stage activity, known as ore-controlling, transportation, and storage, the mineralization in the fracture zone, are a promising area for prospecting.

(2) Regarding the vertical zoning sequence of the primary halo, some boreholes in the No. 21 and No. 25 exploration lines show a superposition of sub-ore and near-ore halos. Distribution of I-1 and I-2 orebodies are parallel to I-3 and I-4, respectively, with wild goose arrangement in space. Considering the spatial superposition characteristics of supra-ore and sub-ore halos (Table 5) reflected by boreholes (ZK2103, ZK2505 and ZK2102), the orebody in the lower part of the known orebody is uncovered. In addition, the similar overprint halo of supra-ore and sub-ore in ZK2107 and Zk2504 or even Zk2505 may indicate the existence of deeply hidden orebodies.

(3) The vertical concentration zoning of the primary halo is the most intuitive method to reveal the distribution of anomalies in the sections. The supra-ore, near-ore, and sub-ore halo elements in two sections well correspond to the orebody areal distribution (Figs. 4 and 5), and the halo scale is consistent with the geochemical characteristics of each element. Besides, the tectonic fracture zone, mineralized alteration zone, and some intrusive bodies can form primary halo anomalies as well.

(4) The two metallogenic stages of exhalative-sedimentation mineralization, and metamorphic-deformation and hydrothermal superimposed mineralization can be well-distinguished in the factor score maps (Figs. 6 and 7). The favorable prospecting area is shown from the F2 principal factor score map, which has a good relationship with the known orebody.

(5) Both the VECC and GP of two cross-sections well reveal the spatial distribution of the orebodies and mineralized alternation zones. Multiple peak values of GP in most of the boreholes (Figs. 8 and 9) indicate that there were several mineralization stages in the study area. To a certain extent, it shows that the metallogenic fluids may have the characteristics of the multiple-periodic and multiple staged activity.

(6) The Trilinear diagram of primary halo (Fig. 10) evidently shows the development and spatial distribution of the primary halos in different elevation segments, corresponding to the different geological bodies, being of great important guidance for the establishment of an ideal model for the superposition of the primary halo.

Based on the identification standards, four prospecting targets are delineated in two profiles (Figs. 11 and 12). The first one is the prediction area (a) (Fig. 11), in which the small I-1 orebody between −300 and 100 m is attached with the I-2 main orebody at the −100 to 400 m elevation interval (Fig. 11). The scale of the primary halo concentration for most elements in this section (from −300 to 0 m) is comparable to

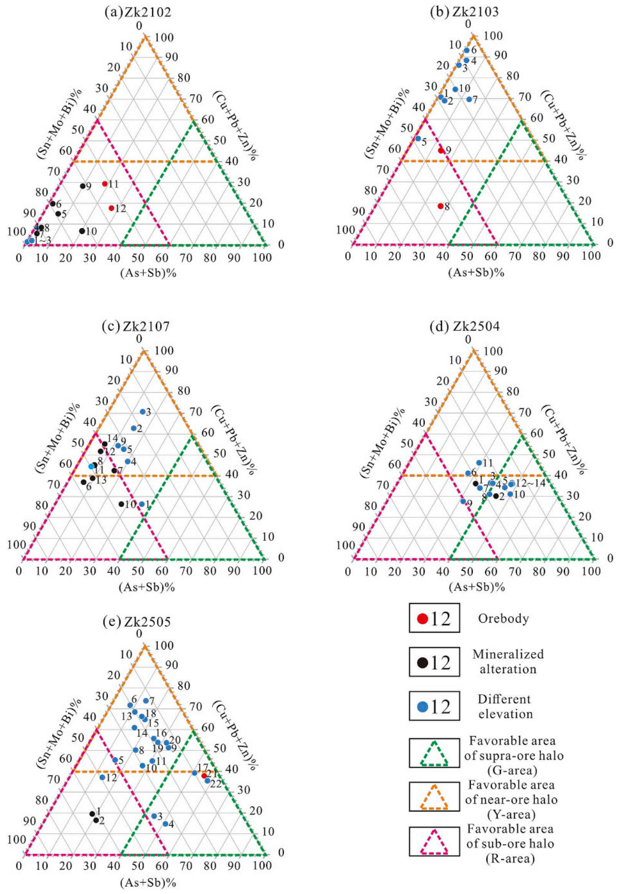


Fig. 10. Trilinear diagrams of the primary halo components for each borehole (Zk2102 (a), Zk2103 (b), Zk2107 (c), Zk2504 (d), Zk2505 (e)). The labels of the points represent different segments of elevation (higher value of the label explains greater depth). Tables S.2(1) to S.2(5) (Supporting material) provide more information about the elevation segments. G-area, favorable area of supra-ore halo; Y-area, favorable area of near-ore halo; R-area, favorable area of sub-ore halo.

that of the upper section (between 0 and 400 m), and even larger than that (e.g. Zn, Cu, Au, As, Sb, Sn, Mo and Bi) of the upper anomaly. In addition, the prediction area (a) with a superposition of sub-ore halo and supra-ore halo at the tail of borehole Zk2107, probably indicating a hidden orebody. From the areal distribution of known orebodies, known as lenticular or stringer assemblages and parallel or attach to each other, in the two sections (Fig. 11), It can be inferred that hidden parallel orebodies with the shape of an en echelon might exist in the prediction area (a). Furthermore, the lithology of this area is consistent with that of the ore-bearing strata. However, the factor score map (Fig. 6) shows no significant elevated values here which may be feasible owing to the excessive sampling distance and the location of the orebodies (I-1 & I-2) between basalt and dacite porphyry. As a result, the high scores were concealed;

The prediction area (b-1) (Fig. 11), located in favorable ore-bearing strata, develops a superposition of sub-ore and supra-ore halos along with the evident primary halo anomalies of many elements (Fig. 4), which can be caused possibly either by I-3 and I-4 orebodies or a hidden orebody. According to the increasing GP (Fig. 5c), there is a high possibility of a hidden orebody or mineralization zone beneath the I-3 and I-4 orebodies. The same favorable conditions mentioned above can also be adapted to No. 25 cross-section (Fig. 12) at the elevation of −400 to −600 m. Considering the extension of orebody (trend from south to north), a slight possibility shows that I-3 and I-4 orebodies may extend to the area (b-2).

The prediction area (c) (Fig. 11) shows a significant primary halo concentration on Bi, Cu, Pb, As elements, accompanying high VECC values at the elevation of 650 m and fluctuating downtrend of GP between 600 and 800 m(Fig. 9). On the other hand, A strong association between Zn, Pb, Au, Ag elements in the F2 score (Fig. 7b), recognized as metallogenic elements assemblage, presents in area (c) (Fig. 11). Furthermore, the area (c) is adjacent to the fracture zone and the lithology is the dacite porphyry rather than ore-bearing strata, which is more possibly correspond to mineralized alternation zone instead of the concealed ore deposit.

Meanwhile, the lithology in prediction area (d) is as same as the one in area (c), including a superimposed sub-ore and supra-ore halos (Table 5) along with vertical geochemical abnormal (Fig. 6). Moreover, a small fracture zone in the dacite porphyry on the lower right side of the prediction area (d), which can also be the reason for these anomalies.

5. Conclusion

The studies we carried out, through analysis of 404 rock samples, demonstrated that the axial primary halo methods (including the axial zoning characteristics, the vertical variation characteristics and the trilinear diagram of primary halo) and statistical method can be applied

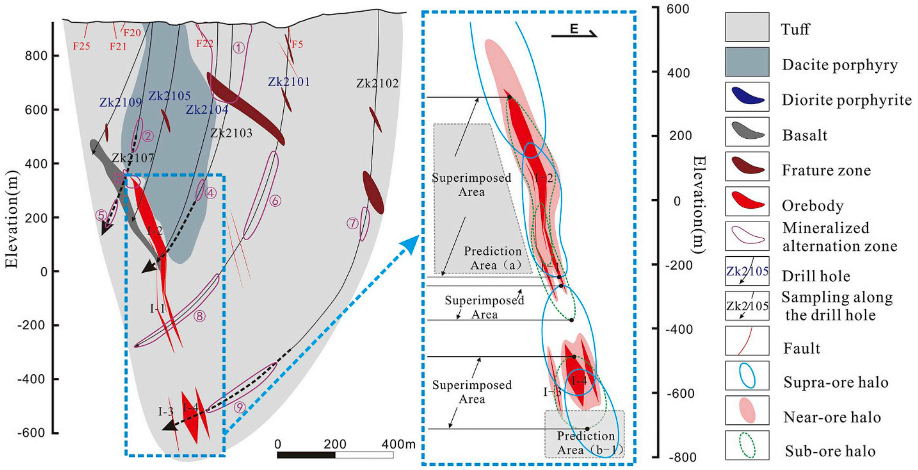


Fig. 11. Vertical superposition model map and deep prediction of primary halo in No. 21 exploration line.

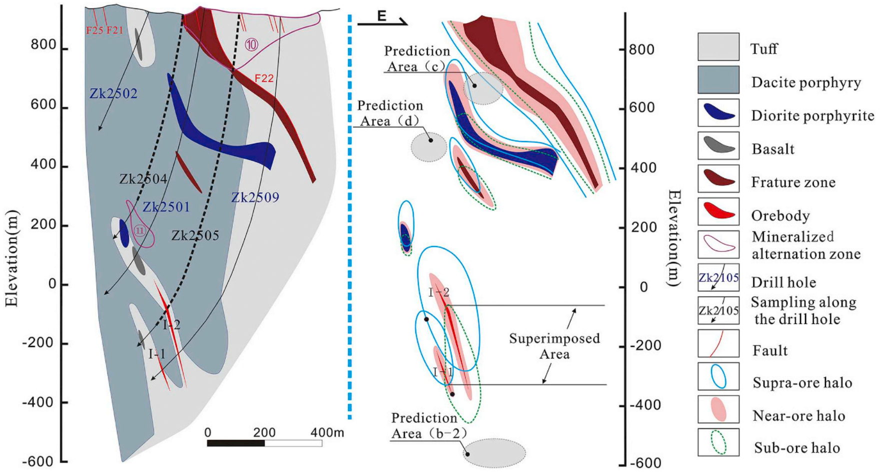


Fig. 12. Vertical superposition model map and deep prediction of primary halo in No. 25 exploration line.

for deep copper-zinc orebody prediction in Ashele mining area. The characteristics of the superimposed haloes of ore deposit are as follows:

(1) Combined with the previous research and the geological condition, it is confirmed that the supra-ore halo element assemblage is As—Sb— Ba—Au, near-ore halo element assemblage is Pb-Zn-Cu-Ag (Au), and sub-ore halo element assemblage is Bi-Co-Mo-Sn. The axial zoning sequence of the primary halo, calculated by the IGZI method, reveals an obvious reverse zoning feature, and most of the reverse zone anomalies can be well explained by the geological bodies.
(2) Considering the axial zoning indexes, the vertical variation characteristic (VECC and GP) and the trilinear diagram of primary halo, when the supra-ore halo is superimposed on the sub-ore halo in the tail of the boreholes, a hidden orebody may lie under the known orebody. The superposition of supra-ore halo and sub-ore halo in the tails of ZK2103, ZK2102 and ZK2505 shows the deep copper-sulfur industrial orebody. Nonetheless, the same characteristics in the tails of ZK2107 and ZK2504 (even in ZK2102 and ZK2505) are feasibly indicators of a hidden orebody at depth. Fluctuations of the geochemical parameter reveal that the deposit has the characteristics of the multiple-periodic and multiple staged overlapping. It can also be concluded from the factor scores: F1 factor score map mainly reveals the later metamorphic-deformation & superimposed hydrothermal mineralization, which the F2 factor score map represents exhalative-sedimentation mineralization.
(3) According to the geochemistry of the axial primary halo of the investigated cross-sections and geology of the mining area, the vertical superposition model maps in Ashele copper-zinc mining area are constructed. Taking the metallogenic potential into account, four promising prospecting areas are delineated.

CRediT authorship contribution statement

Chaojie Zheng: Conceptualization, Methodology, Visualization, Investigation, Writing - original draft, Software, Validation.**Xianrong Luo:**Conceptualization, Methodology, Supervision.**Meilan Wen:**Conceptualization, Methodology, Supervision, Writing - review & editing.**Pooria Ebrahimi:**Writing - original draft, Software, Validation.**Panfeng Liu:**Visualization, Investigation, Writing - review & editing.**Gang Liu:**Visualization, Investigation.**Heping Li:**Data curation.**Chaocheng Wei:**Data curation.**Zesu Chen:**Data curation.**Stefano Albanese:**Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Professor Zuo Renguang for offering the ArcFractal Plugin to apply the C-A fractal model in the study area, and two anonymous reviewer's comments and suggestions, which greatly improve this study. In addition, we are also grateful to the team N.4 of the Xinjiang Geology and Mineral Resources Bureau for the kind support of data collection and especially grateful for the detailed guidance of engineer Xiaogui Wu in the regional geological setting. This study is financially supported by the National Key Research and Development Program of China (Grant Nos. 2016YFC 0600603 & 2017YFC 0602604).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gexplo.2020.106509.

References

Albanese, S., De Vivo, B., Lima, A., Cicchella, D., 2007. Geochemical background and baseline values of toxic elements in stream sediments of Campania region (Italy). J. Geochem. Explor. 93 (1), 21–34. https://doi.org/10.1016/j.gexplo.2006.07.006.
Albanese, S., Vivo, B. de, Lima, A., Cicchella, D., Civitillo, D., Cosenza, A., 2010. Geochemical baselines and risk assessment of the Bagnoli brownfield site coastal sea sediments (Naples, Italy). J. Geochem. Explor. 105 (1–2), 19–33. https://doi.org/10.1016/j.gexplo.2010.01.007.
Albanese, S., Iavazzo, P., Adamo, P., Lima, A., Vivo, B. de, 2013. Assessment of the environmental conditions of the Sarno river basin (south Italy): a stream sediment approach. Environ. Geochem. Health 35 (3), 283–297. https://doi.org/10.1007/s10653-012-9483-x.
Albanese, S., Sadeghi, M., Lima, A., Cicchella, D., Dinelli, E., Valera, P., Falconi, M., Demetriades, A., Vivo, B. de, 2015. GEMAS: cobalt, Cr, Cu and Ni distribution in agricultural and grazing land soil of Europe. J. Geochem. Explor. 154, 81–93. https://doi.org/10.1016/j.gexplo.2015.01.004.
Beus, A.A., Grigorian, S.V., 1977. Geochemical exploration methods for mineral deposits. Earth Sci. Rev. 14 (1), 67–69. https://doi.org/10.1016/0012-8252(78)90045-4.
Carranza, E.J.M., 2012. Primary geochemical characteristics of mineral deposits: implications for exploration. Ore Geol. Rev. 45, 1–4. https://doi.org/10.1016/j.oregeorev.2012.02.002.
Chen, Y., 1996. Ore-forming Conditions and Metallogenic Prognosis of the Ashele Copper-Zinc Metallogenic Belt, Xinjiang, China, 1st ed. Geological Publishing House, Beijing (330 pp).
Chen, Y., Zhao, P., 1998. Zonation in primary halos and geochemical prospecting pattern for the Guilaizhuang gold deposit, eastern China. Nonrenewable Resources 7 (1),

37–44. https://doi.org/10.1007/BF02782507.

Chen, Y., Huang, J., Liang, Z., 2008. Geochemical characteristics and zonation of primary halos of pulang porphyry copper deposit, Northwestern Yunnan Province, Southwestern China. J. China Univ. Geosci. 19 (4), 371–377. https://doi.org/10.1016/S1002-0705(08)60070-9.

Chen, W., Yi, L., Chen, C., Song, Y., Li, G., 2016. Geochemical characteristics of primary halos in No. I Ore Body of Keyue Pb-Zn-Sb-Ag polymetallic deposit, Tibet. Journal of Jilin University(Earth Science Edition) 46 (06), 1711–1723. https://doi.org/10.13278/j.cnki.jjuese.201606110.

Chen, J., Fu, Y., Zhao, J., Tang, Y., 2017. Characteristics of primary geochemical halos and deep prospecting target delineation, Guoluolongwa gold deposit, East Kunlun Orogenic Belt. Geological Science and Technology Information (1), 161–167 + 225.

Cheng, Q., 1999. Spatial and scaling modeling for geochemical anomaly separation. J. Geochem. Explor. 65 (3), 175–194. https://doi.org/10.1016/S0375-6742(99)00028-X.

Cheng, Q., 2003. GeoData Analysis System (GeoDAS) for Mineral Exploration and Environmental Assessment, User's Guide (GeoDAS Phase III). York University, Toronto, Ontario, Canada.

Cheng, Q., Agterberg, F.P., Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. J. Geochem. Explor. 51 (2), 109–130. https://doi.org/10.1016/0375-6742(94)90013-2.

Dai, Z., Qian, Z., Qian, J., Zhang, L., 2017. The features structural superimposed halos of the Lianzigou gold deposit in Xiaoqinling area and its deep prediction. Northwest. Geol. 50 (4), 166–175.

Daya Sagar, B.S., Cheng, Q., Agterberg, F.P., 2018. Handbook of Mathematical Geosciences: Fifty Years of IAMG. xxviii Springer Berlin Heidelberg, New York, NY (914 pàgines).

Field, A.P., 2013. Discovering Statistics Using IBM SPSS Statistics, 4th ed. SAGE, London.

Gao, Z., Fang, W., Fang, R., 2010. Tectonic setting and geochemical characteristics of dacite andesite-basalt volcanic rocks in Ashele copper deposit. Xinjiang. Mineral Deposits (02), 218–228.

Goldberg, I.S., Abramson, G.Y., Los, V.L., 2003. Depletion and enrichment of primary haloes: their importance in the genesis of and exploration for mineral deposits. Geochemistry: Exploration, Environment, Analysis 3 (3), 281–293. https://doi.org/10.1144/1467-7873/03-011.

Gong, Q., Yan, T., Li, J., Zhang, M., Liu, N., 2016. Experimental simulation of element mass transfer and primary halo zone on water-rock interaction. Appl. Geochem. 69, 1–11. https://doi.org/10.1016/j.apgeochem.2016.04.001.

Hannington, M.D., Santaguida, F., ImKjarsgaard, Cathles, L.M., 2003. Regional-scale hydrothermal alteration in the Central Blake River Group, western Abitibi subprovince, Canada: implications for VMS prospectivity. Mineral. Deposita 38 (4), 393–422. https://doi.org/10.1007/s00126-002-0298-z.

Helba, H.A., Khalil, K.I., Mamdouh, M.M., Khalek, I.A.A., 2020. Zonation in primary geochemical haloes for orogenic vein-type gold mineralization in the Quartz Ridge prospect, Sukari gold mine area, Eastern Desert of Egypt. J. Geochem. Explor. 209, 106378. https://doi.org/10.1016/j.gexplo.2019.106378.

Hosseini-Dinani, H., Aftabi, A., 2016. Vertical lithogeochemical halos and zoning vectors at Goushfil Zn–Pb deposit, Irankuh district, southwestern Isfahan, Iran: implications for concealed ore exploration and genetic models. Ore Geol. Rev. 72, 1004–1021. https://doi.org/10.1016/j.oregeorev.2015.09.023.

Konstantinov, M.M., Strujkov, S.F., 1995. Application of indicator halos (signs of ore remobilization) in exploration for blind gold and silver deposits. J. Geochem. Explor. 54 (1), 1–17. https://doi.org/10.1016/0375-6742(95)00003-8.

Li, H., 1998. Primary Halo Model for Buried Ore Prospecting of Large and Super-large Gold Deposits. Metallurgical Industry Press, Beijing.

Li, H., Zhang, W., 1999. Ideal models of overprint of primary halo for large, mega-size blind Au ore deposits. Contributions to Geology and Mineral Resources Research 14 (3), 25–33. https://doi.org/10.6053/j.issn.1001-1412.1999.3.

Li, H., Wang, Z., Li, F., 1995. Ideal models of superimposed primary halos in hydrothermal gold deposits. J. Geochem. Explor. 55 (1), 329–336. https://doi.org/10.1016/0375-6742(94)00063-8.

Li, H., Yu, B., Li, H., Zhang, G., 2010. Structural superimposed halos method for prospecting blind ore-body in the deep of ore-districts. Earth Science Frontiers 17 (01), 287–293. https://doi.org/10.3724/SP.J.1231.2010.06586.

Li, H., Yu, B., Li, H., Zhang, G., 2013. Prediction of blind ore bodies using structural superimposed halo and research methods. Geol. Explor. 49 (1), 154–161.

Li, H., Yu, B., Zhang, D., Zhang, G., 2014. The key technique for applying the structural superimposed halo method to the prediction of the blind ore body. Geophysical and Geochemical Exploration 38 (02), 189–193.

Li, H., Yu, B., Zhang, G., Ma, J., Li, D., 2016a. The development and innovation of blind ore geochemical prospecting method from primary halo and superimposed primary halo to structural superimposed halo. Contributions to Geology and Mineral Resources Research 31 (01), 92–98.

Li, Y., Zhang, D., Dai, L., Wan, G., Hou, B., 2016b. Characteristics of structurally superimposed geochemical haloes at the polymetallic Xiasai silver-lead-zinc ore deposit in Sichuan Province, SW China. J. Geochem. Explor. 169, 100–122. https://doi.org/10.1016/j.gexplo.2016.07.007.

Li, S., Li, Y., Han, Y., Kong, L., 2016c. The characteristics of vertical primary halo zoning in the Louli gold deposit,Handan City,Hebei Province. Geology in China 43 (05), 1637–1644. https://doi.org/10.12029/gc20160513.

Liu, C., Ma, S., 2007. The main achievements of the study on primary halo in the past 50 years in China. Computing Techniques for Geophysical and Geochemical Exploration 29 (S1), 215–221 + 4.

Liu, C., Ma, S., Hu, S., 2008. Petrochemical Characteristic and prognostic indices of volcanogenic hydrothermal ore deposits. Geophysical and Geochemical Exploration (02), 154–158. https://doi.org/10.1016/S1876-3804(08)60015-4.

Liu, C., Ma, S., Hu, Q., 2010. A study of primary halo exploration indexes for metallic ore deposit. Geophysical and Geochemical Exploration 34 (06), 765–771. https://doi.org/10.1016/S1876-3804(11)60004-9.

Liu, C., Hu, S., Ma, S., Tang, L., 2014. Primary geochemical patterns of Donggua Mountain laminar skarn copper deposit in Anhui, China. Geophysical and Geochemical Exploration 139, 152–159. https://doi.org/10.1016/j.gexplo.2013.08.002.

Lopes, R., Betrouni, N., 2009. Fractal and multifractal analysis: a review. Med. Image Anal. 13 (4), 634–649. https://doi.org/10.1016/j.media.2009.05.003.

Luo, X., Wen, M., Ou, Y., 2007. Exploration Geochemistry. Metallurgical Industry Press, Beijing (261 pp).

Pang, X., Zhang, K., Zhang, S., 2016. Metallogenic regularities and deep prospecting of the longmendian silver deposit in Western Henan Province. Bull. Mineral. Petrol. Geochem. 35 (2), 272–278. https://doi.org/10.3969/j.issn.1007-2802.2016.02.006.

Parsapoor, A., Khalili, M., Maghami, M., 2017. Discrimination between mineralized and unmineralized alteration zones using primary geochemical haloes in the Darreh-Zar porphyry copper deposit in Kerman, southeastern Iran. J. Afr. Earth Sci. 132, 109–126. https://doi.org/10.1016/j.jafrearsci.2017.04.027.

Reimann, C., 2008. Statistical Data Analysis Explained: Applied Environmental Statistics With R. Wiley, Chichester (John Wiley [distributor], Hoboken, N.J).

Safronov, N.I., 1936. Dispersion haloes of ore deposits and their use in exploration. ProblemySovetskoyGeologii 4 (13), 41–53.

Shao, Y., 1997. Exploration for Hydrothermal Deposit With Primary Halo Prospecting, 1st ed. iv. Geological Publishing House, Beijing, pp. 143.

Wang, D., 1995. On the Metallogenic Mechanism and Model of the Ashelevolacnogenic Massive Sulfide Copper Deposit, Xinjiang.

Wang, X., Xie, X., Ye, S., 1995. Concepts for geochemical gold exploration based on the abundance and distribution of ultrafine gold. J. Geochem. Explor. 55 (1–3), 93–101. https://doi.org/10.1016/0375-6742(95)00026-7.

Wang, J., Zang, X., Guo, X., Xie, H., 2007. The improved Gregorian's zoning index calculating method. Journal of Jilin University(Earth Science Edition) (05), 884–888.

Wang, Q., Jiang, Y., Yu, H., Li, G., 2011. Primary halo characteristic and concealed ore body prognosis in the Nannong Lead-Zinc Polymetallic, Southwestern Zhejiang Province. Geophysical and Geochemical Exploration 35 (02), 170–175. https://doi.org/10.1007/s12182-011-0118-0.

Wang, C., Carranza, E.J.M., Zhang, S., Zhang, J., Liu, X., Zhang, D., Sun, X., Duan, C., 2013. Characterization of primary geochemical haloes for gold exploration at the Huanxiangwa gold deposit, China. J. Geochem. Explor. 124, 40–58. https://doi.org/10.1016/j.gexplo.2012.07.011.

Wei, J., Liu, C.-q., Zhao, Z.-q., Li, J.-w., 2000. Location prediction of blind ore-bodies in Shuijingtun Gold Mine, Zhangjiakou, China. Journal of Central South University: Science & Technology of Mining and Metallurgy 7 (3), 124–128. https://doi.org/10.1007/s11771-000-0018-8.

Wei, H., Xu, J., Zhao, Z., Jing, Y., Liu, Y., Wu, J., 2011. Analysis of mathematical geology for soil geochemistry in the Caojiba area of Liangcheng County, Inner Mongolia. Geophysical and Geochemical Exploration 47 (03), 473–482. https://doi.org/10.1007/s12182-011-0118-0.

Wu, Y., Yang, F., Liu, F., 2015. ^{40}Ar-^{39}Ar dating of sericite from the brittle ductile shear zone in the Ashele Cu-Zn Ore district, Xinjiang. Acta GeoscienticaSinica 36 (01), 121–126. https://doi.org/10.3975/cagsb.2015.01.15.

Xie, S., Cheng, Q., Ke, X., Bao, Z., Wang, C., 2008. Identification of geochemical anomaly by multifractal analysis. J. China Univ. Geosci. 19 (4), 334–342. https://doi.org/10.1016/S1002-0705(08)60066-7.

Xing, L., Liu, B., 2011. Study of zonation of primary halo of hydrothermal deposits. Acta Geologica Sichuan 31 (04), 489–492 + 495. https://doi.org/10.3969/j.jssn.1006-0995.2011.04.025.

Yang, F., Wang, F., Wang, J., Zheng, K., Liu, F., 2013. LA-MC-ICP-MS U-Pb dating of zircons from volcanic rocks and subvolcanic intrusions in Ashele Cu-Zn ore district, Xinjiang, and its geological implications. Mineral Deposits 32 (05), 869–883. https://doi.org/10.16111/j.0258-2013.05.020.

Yang, F., Wu, Y., Yang, J., Zheng, J., 2016. Metallogenetic model for VMS type polymetallic copper deposits in the Ashele Ore Dense district of Altay, Xinjiang. Geotecton. Metallog. 4, 701–715. https://doi.org/10.16539/j.ddgzyckx.2016.04.006.

Ye, Q., Fu, X., Zhang, X., 1997. Geological characteristics and genesis of the Ashele copper-zinc massive sulfide deposit, Xinjiang. Mineral Deposits (02), 2–11.

Ye, Q., Fu, X., Wang, B., 1998. Metallogeny of polymetallic metallogenic belts on the southern margin of the Altay Mountains, Xinjiang, China. Acta GeologicaSinica 72 (04), 349–357. https://doi.org/10.1088/0256-307X/16/9/027.

Yousefi, M., 2017. Analysis of zoning pattern of geochemical indicators for targeting of porphyry-Cu mineralization: a pixel-based mapping approach. Nat. Resour. Res. 26 (4), 429–441. https://doi.org/10.1007/s11053-017-9334-7.

Yousefi, M., Carranza, E.J.M., 2015. Prediction–area (P–A) plot and C–A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. Comput. Geosci. 79, 69–81. https://doi.org/10.1016/j.cageo.2015.03.007.

Zhang, Y., Gu, X., Chen, W., Dong, S., Jing, L., 2010. The geochemical features of primary halo and the evaluation of deep mineralization prospect of Liubagou gold deposit, Inner Mongolia. Earth Science Frontiers 17 (02), 209–221.

Zhao, L., Liu, W., Shen, Y., Zhan, W., 1989. A model of the primary halo and the geochemical mechanism of halo-ore-forming processes over a strata-bound Pb-Zn-FeS$_2$ deposit. J. Geochem. Explor. 33 (1), 85–97. https://doi.org/10.1016/0375-6742(89)90020-4.

Zhao, J., Chen, S., Zuo, R., Carranza, E.J.M., 2011. Mapping complexity of spatial distribution of faults using fractal and multifractal models: vectoring towards exploration targets. Comput. Geosci. 37 (12), 1958–1966. https://doi.org/10.1016/j.cageo.2011.04.007.

Zheng, Y., Li, D., Zhang, L., Wang, C., Zhang, J., 2015. A preliminary study on the ore metal remobilization of the Ashele VMS Cu-Zn deposit, Xinjiang. Geotecton. Metallog. (3), 542–553. https://doi.org/10.16539/j.ddgzyckx.2015.03.015.

Ziaii, M., Carranza, E.J.M., Ziaei, M., 2011. Application of geochemical zonality coefficients in mineral prospectivity mapping. Comput. Geosci. 37 (12), 1935–1945. https://doi.org/10.1016/j.cageo.2011.05.009.

Zuo, R., 2011. Identifying geochemical anomalies associated with Cu and Pb-Zn skarn mineralization using principal component analysis and spectrum-area fractal modeling in the Gangdese Belt, Tibet (China). J. Geochem. Explor. 111 (1–2), 13–22. https://doi.org/10.1016/j.gexplo.2011.06.012.

Zuo, R., 2012. Exploring the effects of cell size in geochemical mapping. J. Geochem. Explor. 112, 357–367. https://doi.org/10.1016/j.gexplo.2011.11.001.

Zuo, R., Wang, J., 2016. Fractal/multifractal modeling of geochemical data: a review. J. Geochem. Explor. 164, 33–41. https://doi.org/10.1016/j.gexplo.2015.04.010.

Zuo, R., Wang, J., 2019. ArcFractal: an ArcGIS add-in for processing geoscience data using fractal/multifractal models. Nat. Resour. Res. 1–10. https://doi.org/10.1007/s11053-019-09513-5.

Publication No. 5

Ebrahimi, P., & Barbieri, M. (2019). Gadolinium as an emerging microcontaminant in water resources: threats and opportunities. Geosciences, 9(2), 93. https://doi.org/10.3390/geosciences9020093.

geosciences

Review

Gadolinium as an Emerging Microcontaminant in Water Resources: Threats and Opportunities

Pooria Ebrahimi [1,2] **and Maurizio Barbieri** [3,*]

[1] Department of Earth Sciences, College of Sciences, Shiraz University, Shiraz 71454, Iran; pooria.ebrahimi@gmail.com
[2] Department of Earth, Environment and Resources Sciences, University of Naples Federico II 80126, Naples, Italy
[3] Department of Earth Sciences, Sapienza University of Rome, 00185 Rome, Italy
* Correspondence: maurizio.barbieri@uniroma1.it; Tel.: +39-06-4991-4593

Received: 4 January 2019; Accepted: 11 February 2019; Published: 17 February 2019

Abstract: As a result of high doses of paramagnetic gadolinium (Gd) chelates administered in magnetic resonance imaging (MRI) exams, their unmetabolized excretion, and insufficient removal in wastewater treatment plants (WWTPs), large amounts of anthropogenic Gd (Gd_{anth}) are released into surface water. The upward trend of gadolinium-based contrast agent (Gd-CA) administrations is expected to continue growing and consequently higher and higher anthropogenic Gd concentrations are annually recorded in water resources, which can pose a great threat to aquatic organisms and human beings. In addition, the feasibility of Gd retention in patients administered with Gd-CAs repeatedly, and even potentially fatal diseases, including nephrogenic systemic fibrosis (NSF), due to trace amounts of Gd have recently arisen severe health concerns. Thus, there is a need to investigate probable adverse health effects of currently marketed Gd-CAs meticulously and to modify the actual approach in using Gd contrast media in daily practice in order to minimize unknown possible health risks. Furthermore, the employment of enhanced wastewater treatment processes that are capable of removing the stable contrast agents, and the evaluation of the ecotoxicity of Gd chelates and human exposure to these emerging contaminants through dermal and ingestion pathways deserve more attention. On the other hand, point source releases of anthropogenic Gd into the aquatic environment presents the opportunity to assess surface water—groundwater interactions and trace the fate of wastewater plume as a proxy for the potential presence of other microcontaminants associated with treated wastewater in freshwater and marine systems.

Keywords: water contamination; adverse health effects; ecotoxicity; surface water-groundwater interaction; wastewater treatment; magnetic resonance imaging (MRI); gadolinium-based contrast agents (Gd-CAs)

1. Introduction

The upward trend of mining and using lanthanides (LAs: La to Lu) in the modern world has been accompanied by the release of similarly increasing amounts of those chemicals in the environment, where their concentrations can exceed their natural values in orders of magnitude [1]. According to articles published over the last three decades, amongst LAs, anthropogenic Gd (the highly stable MRI Gd compounds that have not been fully removed in WWTPs and have been transferred to surface water) seems to be the most ubiquitous water contaminant in many countries, including Australia [2], the Czech Republic [3], Italy [4,5], France [6], the USA [7], Germany [7–9], England [10], Japan [11,12], and South Korea [13]. However, no anomalous Gd concentrations in some urban rivers, such as the Chao Phraya that runs through a densely populated area in Thailand [14], implies that

Gd_{anth} inputs are mainly restricted to regions with a highly developed healthcare system and mass application of MRI tests. Positive Gd anomalies of variable size have been observed in the influent and effluent of WWTPs [13,15–17], rivers [2,7,9–12,18,19], seawaters [1,12,15,20], groundwaters [21], and tap waters [7,10] due to anthropogenic Gd input.

The discharge of treated wastewater containing Gd_{anth} into nearby surface water bodies has provided some opportunities for hydrological studies and tracing of different microcontaminants in the treated wastewater plume [2,3]. On the other hand, it has been accompanied with some community concern, which is not irrelevant to the increasing prevalence of published scientific articles reporting the occurrence of micropollutants originating from WWTP effluents in the environment [22,23]. Microcontaminants of emerging concern (MECs), which include Gd_{anth}, are pollutants that can be detected using analytical equipment with quite low detection limits. The mechanisms of diffusion-driven transport/uptake and active transport/bioaccumulation of the Gd chelates by living organisms are of great importance, particularly when investigating feasible adverse consequences for human health if these compounds reach the human food chain. However, due to a lack of regular monitoring schemes for MECs, their toxicity is poorly understood and they are not currently regulated in the human body and the natural environment [24,25].

The terms of rare elements (REs) or rare earth elements (REEs) regularly used in the literature have been confusing since they refer to different chemical elements depending on the author. REEs are commonly related to the lanthanides comprising of 15 elements, including lanthanum (La), cerium (Ce), praseodymium (Pr), neodymium (Nd), promethium (Pm), samarium (Sm), europium (Eu), gadolinium (Gd), terbium (Tb), dysprosium (Dy), holmium (Ho), erbium (Er), thulium (Tm), ytterbium (Yb), and lutetium (Lu), in the periodic table of elements [26,27]. Amongst the lanthanides, only promethium (Pm), a man-made radioactive element, has not yet been found in the natural environment; however, the other 14 elements are relatively abundant in geological formations and soils in a variety of chemical forms [28]. In another classification, the term, REEs, refers to the lanthanides, yttrium (Y), and scandium (Sc), which are regularly found in the same mineral assemblages due to their almost identical physicochemical characteristics [26,29]. In this review, as recommended by Kabata-Pendias [27], the term, lanthanides (LAs), will be used for the elements from La to Lu in order to avoid any confusion.

Although there is a decreasing trend in the abundance of LAs across the lanthanide series, a greater abundance of elements with an even atomic number described by the Oddo-Harkins rule leads to a zigzag pattern in the LA plots [30,31]. The elements have been categorized into different subgroups by different researchers and it continues to develop on the basis of recent scientific breakthroughs and specific aims [29]. Considering articles relating to water resources, LAs are either grouped into three classes, namely light LAs (LLAs), medium LAs (MLAs), and heavy LAs (HLAs) [32,33], or only two subgroups, including LLAs and HLAs [13,34,35]. Across the lanthanides, there is a difference between the organic complexation behavior of LLAs (La to Eu) and HLAs (Tb to Lu). At the junction of the two series of elements, Gd behaves as an HLA for the outer sphere (weak) complexation and as an LLA for the inner sphere (strong) complexation [36]. Although the majority of the LAs mainly exist in the 3+ state, Nd (2+, 3+, and 5+) together with Dy (2+, 3+, and 4+) exhibit variable valences [28].

To date, no organized monitoring schedule for Gd complexes has resulted in a comprehensive understanding of their toxicity and regulation in the body and the environment. Hence, this review introduces anthropogenic Gd as an overlooked emerging microcontaminant that poses a threat to human and ecological health. The main objectives of the current article review are to understand: (1) sources and fate of Gd_{anth} in the environment; (2) challenges versus opportunities considering Gd_{anth} content in water resources; and (3) future research directions to fill the gaps in our understanding.

2. Gd Discovery, Its Properties, Global Production, and Use

Jean-Charles Gallisard de Marignac investigated the mineral, samarskite, in 1880 and found a new element, which was later named gadolinium to emphasize the great importance of the original

REE mineral, gadolinite. Natural Gd is a mixture of six stable isotopes, which can be arranged in the abundance order of ^{158}Gd (24.84%) > ^{160}Gd (21.86%) > ^{156}Gd (20.47%) > ^{157}Gd (15.65%) > ^{155}Gd (14.80%) > ^{154}Gd (2.18%), and a radioisotope, ^{152}Gd (0.20%). This metal is a silvery-white, ductile and malleable LA, with a white oxide and colorless salts. Whilst it tarnishes in moist air and reacts slowly with water, it is relatively stable in dry air [37].

Over the last half-century, there was an exponential growth in the world production of rare-earth oxides (lanthanides, Sc, and Y) and the figure reached about 130,000 metric tons by 2015 because of the intense technological breakthroughs (Figure 1). Pure LAs were first prepared in 1931, but the use of individual LAs was negligible until the improvement of separation and metallurgical technologies in the 1950s [38]. Approximately four decades later, Bau and Dulski [7] published the first scientific article about the anthropogenic anomaly of Gd in the environment (river water) corresponding to the increasing use of Gd-CAs in MRI tests (considering Figure 1 and Figure 4, from the 1990s onward, the global rare-earth oxide production has been comparable to the number of MRI exams per 1000 inhabitants). Global reserves of rare-earth oxides (REOs) were estimated to be 100 million tonnes in 2000, with China possessing the largest proportion (43%) of those reserves [39]. In terms of REE global production, China outnumbered other countries and supplied 98% of the REE demand [40]. Regarding their unique properties and unparalleled performance, they are critical resource materials for high technology industries and cannot currently be replaced by alternative materials [41,42].

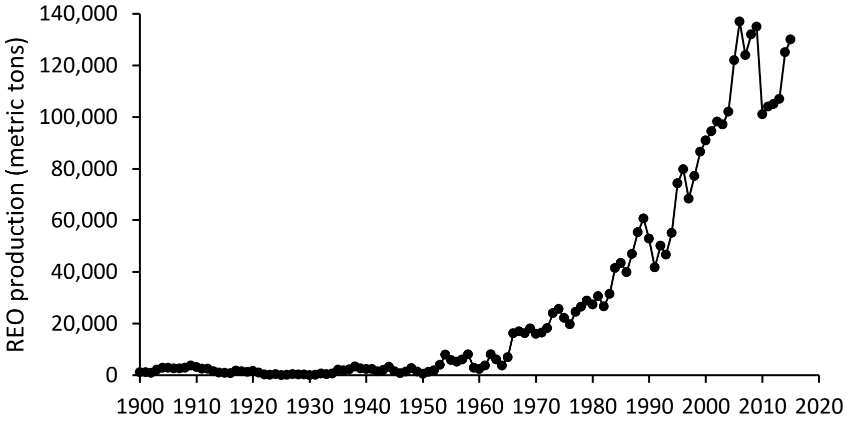


Figure 1. Global rare-earth oxide (REO) production [43].

The end uses of LAs can be divided into two broad categories, namely 'process enablers' and 'product enablers'. In the former, LAs are used in the process of production and not in the end product. In the latter, they give unique characteristics to advanced materials and play a crucial role in the performance of high-technology products [44]. The main applications of Gd in high-tech industries are listed in Table 1.

Table 1. Different high-technology applications of Gd.

Application Area	Description	Reference
Glass industry	Gd, Ce, Pr, Nd, and Er are utilized in the manufacture and polishing of glass products, including lenses.	[26]
Electronics	(1) Gd and Eu have been used for various phosphors in computer monitors and color television tubes; (2) regarding the unique magnetic characteristics of Gd, the element is used in magneto-optic recording technology to handle computer data; (3) gadolinium yttrium garnets are used in microwave ovens to produce the microwaves; (4) at room temperature, Gd displays a strong magnetocaloric effect and can be efficiently used for either magnetic refrigeration or testing other magnetic refrigeration devices; (5) permanent magnets containing Gd, Nd, Dy, and Tb are used in numerous electrical and electronic components, and generators for wind turbines; (6) Gd-containing crystals are used in high-power pulsed lasers; and (7) Gd oxide has been used in the production of optical fibers.	[26,45–48]
Medicine	Gd-CAs are used as paramagnetic markers and contrast agents in MRI for medical diagnoses, including vascular, myocardial, orthopedic, oncologic, inflammatory, and neurological diseases.	[49–55]
The airline and shipbuilding industries	In the airline and shipbuilding industries, neutron radiography is used to search for hidden flaws and structural weakness in hulls and fuselages. In addition to the usual film required for X rays, an additional screen (a conversion screen), which is mostly made of Gd, is placed in close contact with the film.	[56]
Car industry	Ceriumoxide and a minor fraction of Gd are generally added to three-way catalysts to promote the watergas shift reaction.	[57]
Metallurgy	The addition of Gd to Fe and Cr alloys improves their workability and resistance to oxidation and high temperatures.	[58]
Energy	The high melting points and high thermal neutron absorption cross-sections of Gd, Sm, Eu, and Dy are two properties that make them ideal for uses in nuclear reactor control rod applications. The rods are raised out of or lowered into the reactor, allowing for more or fewer neutrons to remain in the nuclear fission reaction. Power plant radiation leaks can also be detected by Gd, being the most efficient detector.	[26,45,58]

3. Natural Gd

Monazite, bastnaesite, and gadolinite account for Gd typical minerals, and feldspars, apatite, allanite, sphene, fluorite, and zircon are the possible host minerals [59]. Whilst the data on individual LAs in soil are scanty, the geological origin of parent rocks and their mineral composition evidently determine the trend of LA distribution in various soils [28]. According to Korzh [60], the rate of Gd deposition from oceans via the atmosphere onto soils is low, being 1 to 10 t/yr. On average, cabbages from the background area of Denmark contain Gd within the range of 0.15–2.60 µg/kg wet weight [61]. Eriksson [62] reported 0.001 to 0.01 µg/L Gd in rainwaters collected in Sweden during 1999.

Data for the LAs are among the most widely used geochemical tools in the earth and environmental sciences [13,19,33,63,64]. In unpolluted natural fresh waters, LA concentrations commonly range between ng/g and low-pg/g (ppb and ppt), and saline waters are typically in the fg/g (ppq) range [65]. Total Gd concentrations and Gd anomalies found in uncontaminated aqueous samples are provided in Table 2. Chemical weathering is a prominent process that partially mobilizes lanthanides as dissolved loads in surface water, during which physicochemical characteristics, such as pH, salinity, flow conditions [66–69], and bedrock lithology, regulate the dissolved LA content [69–71]. LA patterns are thought to be great indicators of overall river basin processes (e.g., chemical weathering and transportation processes) due to the relative consistency of physicochemical properties and bedrock types within a river; except in cases where either atmospheric wet and dry deposition or anthropogenic lanthanides significantly alters the dissolved LA load by orders of magnitude [19,72–74]. Due to this invariable coherent behavior in biogeochemical processes, LA patterns are a prevalent geochemical tool for the interpretation of the dissolved load content of natural waters [24,75–77].

Table 2. Total Gd concentrations and Gd anomalies reported in uncontaminated aqueous samples. The subscript, SN, denotes the normalization to the Gd content in the continental crust (regularly PAAS was used) and the superscript * denotes the natural background value.

Sample	Study Area	Filtration	Acidification	Instrument	Detection Limit	Gd	Gd_{SN}/Gd^{*}_{SN}	Reference
Tapwater	Eastern Berlin, Germany	0.2 μm	pH = 1.8−2.0	ICP-MS	-	2.3–4.2 pmol/L	0.95–1.70 [1]	[10]
	Wivenhoe Dam, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	16.0 pmol/kg	1.07 [2]	[16]
	Somerset Dam, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	64.3 pmol/kg	1.09 [2]	[16]
	Hinze Dam, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	233.1 pmol/kg	1.08 [2]	[16]
	North Pine Dam, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	11.1 pmol/kg	1.07 [2]	[16]
	Mt Crosby Weir, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	21.9 pmol/kg	1.05 [2]	[16]
	Brisbane, Australia	0.22 μm	pH = 1.5	ICP-MS	480 fmol/L	15.5 pmol/kg	1.04 [2]	[16]
	Prague, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	2.1 nmol/m^{3}	1.20 [2]	[3]
River water	Wangsuk stream, South Korea	0.2 μm	pH = 2.0	Q-ICP-MS	-	26.0 pmol/L	1.34 [3]	[13]
	Anyang stream, South Korea	0.2 μm	pH = 2.0	Q-ICP-MS	-	26.0 pmol/L	1.21 [3]	[13]
	Changneung stream, South Korea	0.2 μm	pH = 2.0	Q-ICP-MS	-	35.6 pmol/L	1.39 [3]	[13]
	Wiembach Creek, Germany	0.2 μm	pH = 1.8–2.0	ICP-MS	-	1.5 ng/kg	0.98 [1]	[9]
	Spring Creek, USA	0.2 μm	pH = 2.0	ICP-MS	0.01–0.1 pg/mL	26.3 pmol/L	1.15 [3]	[18]
	Delaware River, USA	0.2 μm	pH = 2.1	ICP-MS	0.01–0.1 pg/mL	23.9 pmol/L	1.47 [3]	[18]
	Ibi River, Japan	0.45 μm	pH = 1	ICP-MS	0.027 ng/L	3.5 ng/L	1.36 [3]	[12]
	Nagara River, Japan	0.45 μm	pH = 1	ICP-MS	0.027 ng/L	3.9 ng/L	1.41 [3]	[12]
	Kiso River, Japan	0.45 μm	pH = 1	ICP-MS	0.027 ng/L	5.3 ng/L	1.31 [3]	[12]
	Amazon River, South America	0.2 μm	-	ICP-MS	-	0.0123 μg/L	1.20 [1]	[78]
	Toshibetsu River, Japan	0.2 μm	pH = 1.9–2.1	ICP-MS	-	77.4 pmol/kg	1.20 [3]	[7]
	Västerdalälven River, Sweden	0.2 μm	pH = 1.9–2.1	ICP-MS	-	265.0 pmol/kg	1.20 [3]	[7]
	Dhünn River, Germany	0.2 μm	pH = 1.9–2.1	ICP-MS	-	8.9 pmol/kg	1.60 [3]	[7]
Groundwater	Hind Well, Hawaii	0.45 μm	pH < 2	HR-ICP-MS	0.5–6.0 pmol/kg	41.2 pmol/kg	1.08 [3]	[21]
	Berlin, Germany	0.2 μm	pH = 1.8–2.0	ICP-MS	-	8.4 pmol/L	1.12 [1]	[10]
	Vlastejovice spring, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	19.4 nmol/m^{3}	1.00 [2]	[3]
Lake water	Lake Värmeln, Sweden	0.2 μm	pH = 1.9–2.1	ICP-MS	-	232.0 pmol/kg	1.20 [3]	[7]
Seawater	North Atlantic seawater	0.45 μm	pH = 1.8–2.0	ICP-MS	-	5.4 pmol/L	1.60 [4]	[15]
	Japan Sea (Nie coast)	0.45 μm	pH = 1	ICP-MS	0.027 ng/L	1.9 ng/L	0.96 [3]	[12]
	Western North Pacific, Japan	0.04 μm	pH < 1.5	ICP-MS	-	2.1 pmol/kg	1.05 [3]	[79]

ICP-MS = Inductively coupled plasma-mass spectrometry; Q-ICP-MS = Quadrupole-inductively coupled plasma-mass spectrometry; HR-ICP-MS = High-resolution-inductively coupled plasma-mass spectrometry. [1] Calculated using linear regression method; [2] Calculated from a third order polynomial fit; [3] $Gd_{SN}/Gd^{*}_{SN} = Gd_{SN}/(0.33Sm_{SN} + 0.67Tb_{SN})$; [4] $Gd_{SN}/Gd^{*}_{SN} = Gd_{SN}/10\hat{}(2\log Sm - \log Nd)$.

After excluding the likely anomalous LAs (La, Ce, Eu, Gd, and Lu), various methods can be used to calculate natural background Gd [3,15,16,80], including (i) linear (or geometric) extrapolation from either the HLAs or the LLAs, (ii) linear (or geometric) interpolation between a HLA and a LLA, or (iii) the third order polynomial fit, which models the shape of the normalized LA pattern. When employing LA values to calculate Gd anomalies (for example, $Gd_{SN}/Gd^{*}_{SN} = Gd_{SN}/(0.33Sm_{SN} + 0.67Tb_{SN})$), the use of arithmetic means is not correct because the logarithmic scale of LA plots leads to serious errors (particularly for steep normalized LA patterns) [81]. Amongst the abovementioned approaches, the last one is the only method that makes no implicit assumption whether Gd behaves as an LLA or an HLA. The application of this method is consequently more compelling as it can be employed for the whole range of waters, from wastewater to freshwater and seawater. Further, with regards to the results obtained by this approach, comparing results between different studies and geographical areas can be facilitated [20].

To detect LA anomalies in the sample of interest and improve their comparison with data from elsewhere, the naturally occurring Oddo-Harkins effect (the saw-tooth effect) in the lanthanide sequence needs to be smoothed out by dividing each measured LA value by the concentration of the same element in the continental crust. To date, LA contents of the Post-Archean Australian Shale (PAAS; [81]), the North American Shale Composite (NASC; [82]), the upper continental crust (UCC; [83]), and the Average Shale data [84] have been used to normalize LA concentrations in water [13,64,85,86]. PAAS, the most widely used composite as it represents the best average of the earth's crust [3], has been evaluated for abundances of the monoisotopic LAs and no accuracy was found in terms of Pr and Tm [87]. Thus, Lawrence et al. [88] used the Mud of Queensland (MUQ; [87], a local composite sediment average, for the normalization and calculation of Gd anomalies. Figure 2 represents MUQ-normalized LA patterns of river waters with and without anthropogenic Gd and wastewater samples with a remarkable positive Gd anomaly. In the typical LA patterns for natural water, the relative concentrations of HLAs are higher than those of LLAs [12]. Moreover, there is a continuous trend from La to Gd prior to a decline to a lower level and another trend from Tb to Lu in the LA plot of pristine rivers, such as Västerdaläven and Toshibetsu and Dhünn [7] (Figure 2a). The larger slopes from LLAs to HLAs might be due to the higher complexing abilities of HLAs with dissolved organic substances and the preferential incorporation of LLAs in the formation of secondary minerals [12,24,89]. However, in waters with less organic substances, submicron particles or colloids containing LLAs more than HLAs pass through the membrane filters with a pore size of 0.45 µm and lead to a flat LA distribution pattern [89,90].

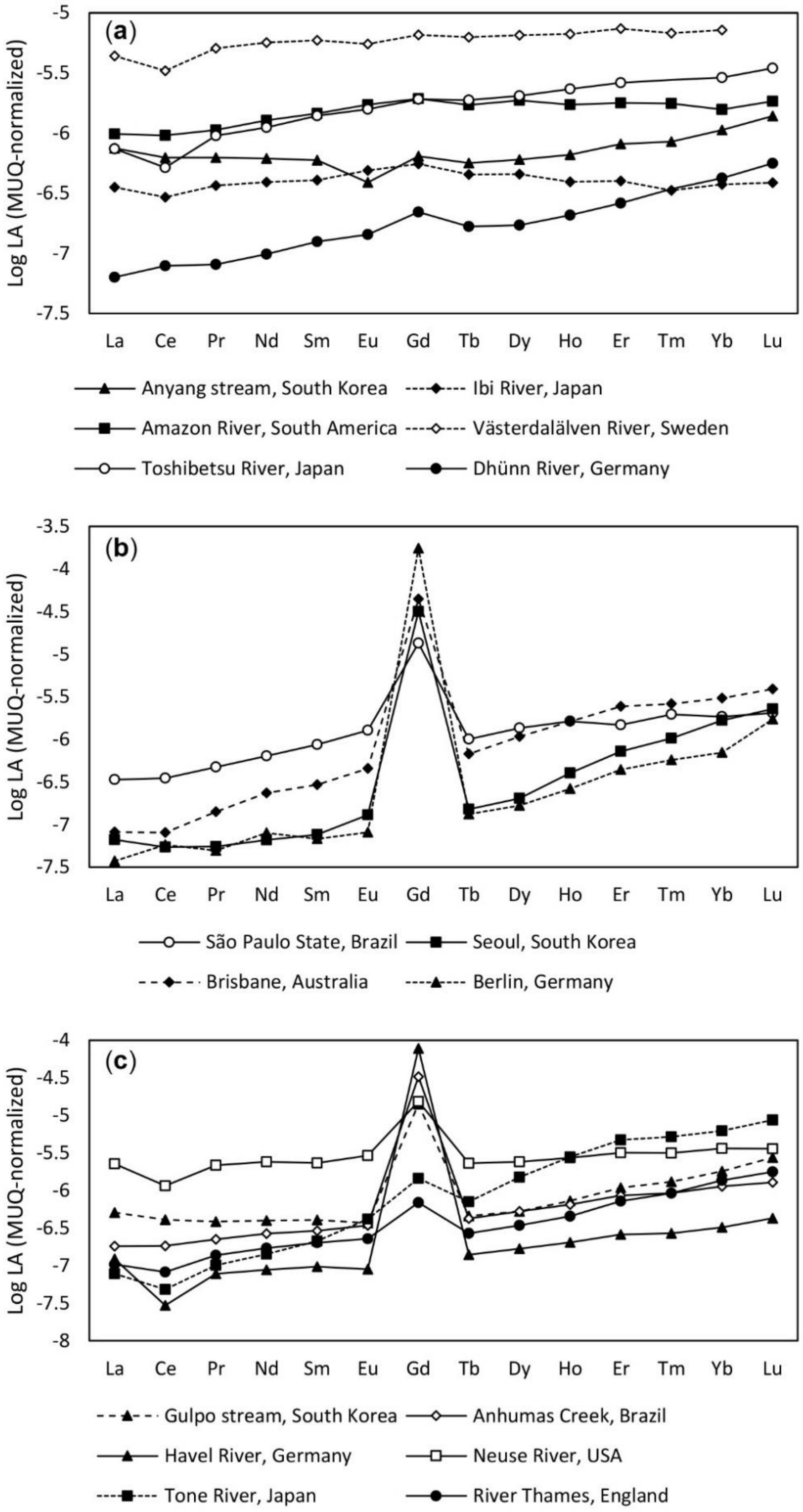


Figure 2. MUQ-normalized LA patterns of: (**a**) River waters without Gd_{anth}; (**b**) WWTP effluents with pronounced positive Gd anomaly; (**c**) river waters with Gd_{anth}. The data are from Bau and Dulski, [7], Nozaki et al. [11], Gaillardet et al. [78], Zhu et al. [12], Lawrence et al. [16], Kulaksız and Bau [10], de Campos and Enzweiler [17], Song et al. [13], and Smith and Liu [19].

The calculation of Gd anomaly using a formula is based on the assumption that normalized values of LAs vary smoothly with an increasing atomic number. Although no positive Gd anomaly is expected in pristine river waters, Bau and Dulski [7] reported Gd_{SN}/Gd^{*}_{SN} ratios of 1.2 for the Västerdaläven and Toshibetsu rivers and 1.9 for the Dhünn river. Since there was a smooth trend from Pr to Gd, and another smooth trend from Tb to Lu in normalized LA patterns of the collected river water samples (Figure 2a), Bau and Dulski [7] concluded that Gd_{SN}/Gd^{*}_{SN} ratios above the unity do not necessarily indicate an occurrence of positive anomalous Gd and the complete normalized LA pattern should also be considered. Due to the unique position of Gd amongst the two groups with different complexation behavior, anomalous Gd may appear regarding the complexation behavior of its immediate neighbors as well [36].

Slight, positive Gd anomalies were identified as a common feature of seawater, which systematically changes with depth. The observed rapid changes in the upper several hundred meters suggest that both ocean circulation and biogeochemical cycles greatly govern seawater Gd distribution [79]. Irrespective of the normalization problem, natural positive Gd anomaly in seawater is also attributed to the "tetrad" effect because the considerable high Gd stability might result from the half-filled 4f electron shell [91]. Further, positive Gd anomaly and distinct negative Tb and Tm anomalies in shale-normalized LA patterns of seawaters were explained by the complexation behavior of LAs with carbonate ions in seawater [92].

4. Anthropogenic Gd (Gd_{anth})

Contrary to the highly stable Gd-CAs used in medicine, the Gd load emitted by other sources or the Gd used in other high-technology applications (Table 1) probably have different toxicities and responses to treatment processes. It seems unlikely that the Gd utilized for these purposes cannot be removed during sewage treatment and end up polluting water resources. The figures for patients who undergo MRI exams and the global administration of Gd-CAs will continue to grow because of the increasing number of available MRI scanners worldwide. Therefore, a large and continuously increasing amount of Gd enters the aqueous environment and enriches Gd over LAs more and more [93]. However, there is no standard for Gd-CAs that regulates their concentrations in sewage effluents or receiving water bodies, partly because of the insufficient knowledge about the ecological effects of LAs [1]. Regarding the stability of Gd-CAs and their potentially high half-lives (i.e., more than three months [94]) in the environment, it is of the utmost importance that the amount of Gd_{anth} in the aquatic environment and also its environmental fate are investigated in more detail. Generally, the inputs and concentrations of Gd_{anth} into the public sewage system and the environment vary greatly depending on the population size, level of access to health systems, and the number of hospitals and clinics performing MRI imaging [1]. In the following sections, elevated concentrations of Gd in water resources, which can be traced back to the application of Gd-CAs in MRI exams and the inefficiency of treatment mechanisms in WWTPs, are reviewed to highlight the potential impacts and benefits of mitigation measures.

4.1. Gd-CAs in MRI

Since the first generation of an image using nuclear magnetic resonance in 1973, non-invasive diagnostic techniques, such as MRI, have been developed as an indispensable method in modern medicine. Although contrast agents are not necessary for MRI exams, 0.1 mmol of Gd per kg of body weight (average body weight is 70 kg [95]) has been applied in about 40% of MRI exams [96,97]. In comparison with other paramagnetic elements, such as manganese, dysprosium, iron, and copper, the paramagnetic characteristics of Gd^{3+} are unparalleled [98]. The seven unpaired electrons of Gd^{3+} account for its strong magnetic momentum, enabling the element to polarize ^{1}H water protons, which indicate an increased T1 relaxation rate when aligned in magnetic fields [99]. It leads to higher signal-to-noise ratios (SNRs), clarity, contrast, and detail for the differentiation of tissues in the MRI images. Since clinical approval of the first contrast agent (Magnevist®) in 1987, the overall administration of various Gd chelates is estimated to have been more than 300 million [100]. Today, the paramagnetic Gd chelates are annually used in about 30 million MRI exams, which is approximately 50% greater than the estimated 2005 figure [100,101].

Regarding the chelate type and charge, Gd-CAs can be divided into linear/macrocyclic and ionic/non-ionic groups. A wide range of Gd-CAs are commercially available and their stability is dependent on the conditional thermodynamic stability constant (K_{cond}), thermodynamic stability constant (K_{therm}), and kinetic stability (Figure 3; Table 3) [97]. Higher values of K_{cond} (at physiological pH tests) and K_{therm} imply a greater Gd affinity for ligands, higher stability of Gd complexes, and less free gadolinium ion [102,103]. The rate of Gd release from its ligand (half-life) is also tested at a low pH using kinetic stability. Greater kinetic stability denotes a longer period of time prior to dissociation (Table 3). Ionic Gd-CAs are more stable than nonionic ones and the stability of macrocyclic compounds is higher than linear compounds. Hence, ionic macrocyclic agents are the most stable Gd chelates. Macrocyclic molecules bind strongly to Gd in an organized rigid ring; however, linear nonionic Gd-CAs have open chains and weaker binding to Gd. Compared to linear agents, macrocyclic agents are more stable in vivo. Low-stability Gd-CAs (linear, nonionic compounds) likely undergo transmetallation, release free Gd that deposits in tissues, attract fibrocytes, and therefore initiate the process of fibrosis [101,104].

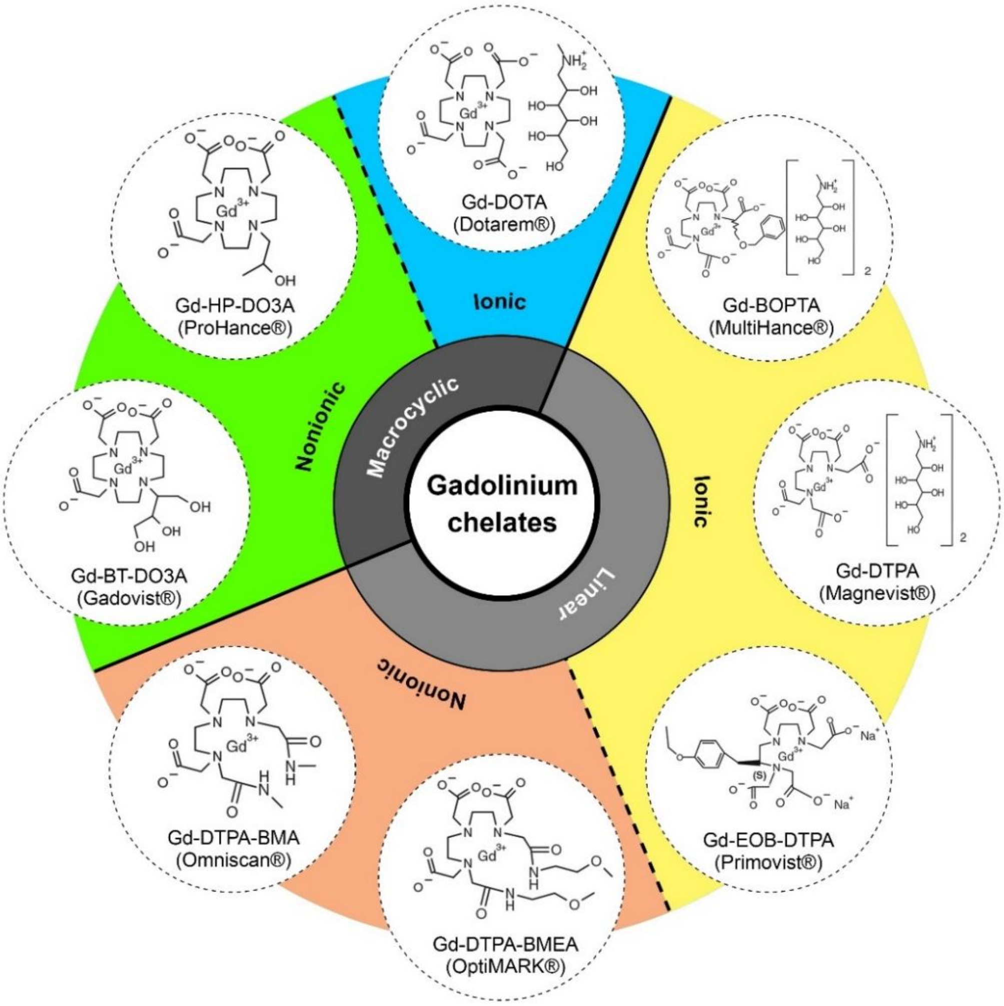


Figure 3. Structure of currently marketed Gd-CAs used for MRI [102].

Table 3. Characteristics of currently marketed Gd-CAs approved by the U.S. Food and Drug Administration for body magnetic resonance imaging [101,103,105–110] and the proportion of pediatric radiology departments used Gd-CAs in 2010 and 2016 [111,112]. Dosage of all Gd-CAs is 0.1 mmol/kg, with the exception of gadoxetate disodium (0.025 mmol/kg).

Chemical Structure	**Chemical Name (Acronym)**	**Trade Name**	**Manufacturer**	**Conditional Thermodynamic Stability Constant (Log K_{cond} at pH 7.4)**	**Thermodynamic Stability Constant (Log K_{therm})**	**Kinetic Stability T1/2 at pH = 1.0 and T = 25 °C**	**NFS Risk [1]**	**Global Administration in 2010 (million)**	**% of Clinics Used Each Gd-CA in 2010**	**% of Clinics Used Each Gd-CA in 2016**
Linear nonionic	Gadodiamide (Gd-DTPA-BMA)	Omniscan®	GE Healthcare (Little Chalfont, UK)	14.9	16.9	<5 s	High	49	10	0
Linear nonionic	Gadoversetamide (Gd-DTPA-BMEA)	OptiMARK®	Guerbet (Villepinte, France)	15	16.6	<5 s	High	3.5	2	0
Linear ionic	Gadopentetate dimeglumine (Gd-DTPA)	Magnevist®	Bayer HealthCare (Berlin, Germany)	17.7	22.1	<5 s	High	105	81	15
Linear ionic	Gadobenate dimeglumine (Gd-BOPTA)	MultiHance®	Bracco Diagnostics (Milan, Italy)	18.4	22.6	<5 s	Intermediate	7.5	38	9
Linear ionic	Gadoxetate disodium (Gd-EOB-DTPA)	Eovist®/Primovist®	Bayer HealthCare (Berlin, Germany)	18.7	23.5	N/A [2]	Intermediate	0.4	-	-
Macrocyclic nonionic	Gadoteridol (Gd-HP-DO3A)	ProHance®	Bracco Diagnostics (Milan, Italy)	17.1	23.8	3.9 h	Low	15	14	4
Macrocyclic nonionic	Gadobutrol (Gd-BT-DO3A)	Gadovist®	Bayer HealthCare (Berlin, Germany)	14.7	21.8	43 h	Low	6	39	42
Macrocyclic ionic	Gadoterate meglumine (Gd-DOTA)	Dotarem®	Guerbet (Villepinte, France)	19.3	25.6	338 h	Low	22.4	N/A	38

[1] [113] [2] not applicable.

Polyaminocarboxylic acid chelating agents are mostly complexed with Gd to produce thermodynamically stable linear or macrocyclic complexes, guaranteeing their fast, complete, and unmetabolized urinary excretion from the human body, and avoiding the high toxicity of free Gd^{3+} [101,106,114,115]. Hence, Gd-CAs pose a lower risk of allergies and interaction with patients than other contrasts, such as iodinated contrast [96].

4.2. Gd Determination in the Aqueous Environment

As indirect potable reuse of purified wastewater poses a threat to wildlife and human health, it is crucial to monitor micropollutants in the environment. However, analytical detection limits restrict our capability to quantify their presence and assess their human health consequences. Therefore, it is imperative to be equipped with some alternative strategies to overcome the challenges [16].

Analytically, the most accurate and precise techniques are isotope dilution methods requiring the equilibration of an enriched isotopic spike with the sample before analysis. For this reason, obtaining the concentration of the monoisotopic elements (Y, Pr, Tb, Ho, and Tm) by isotope dilution is impossible. The importance of the monoisotopic element concentrations for quantifying the behavior of the 'anomalous' elements was enlightened when more LA data for seawater and hydrogenous sediments became available [116]. Utilizing multi-element isotope dilution inductively coupled plasma-mass spectrometry (ID ICP-MS), Behrens et al. [117] reported the first application of the automated seaFAST-pico system in the offline mode for rapid and robust purification, pre-concentration, and analysis of dissolved LAs from small volumes (11 to 12 ml) of seawater. In extraction techniques (either solid phase extraction or liquid—liquid extraction), samples are exposed to other reagents, making it quite challenging to maintain process blanks below the levels required to quantify the LA accurately [118]. Nonetheless, using external calibration, ICP-MS methods are capable of quantifying the complete set of LAs and produce a dataset that may be used to elucidate the origin and significance of the LA pattern. When ICP-MS is used to analyze fresh and marine waters directly, sample manipulation and the corresponding risk of blank contamination are minimized, leading to even better quality data [65].

Lawrence et al. [16] demonstrated an ICP-MS technique that was capable of measuring Gd_{anth} in tap waters, the surface waters of water supply reservoirs, and WWTP effluents at concentrations as low as 48 fmol/L without the need for preconcentration. It was roughly six orders of magnitude lower than the detection limits for chemical methods determining representative pharmaceutical compounds in natural water resources (for instance, Diazepam in river water) where the limit of detection is 0.02 μg/L (70 nmol/L) after a 1000-fold preconcentration [119]. They collected filtered water samples in trace metal cleaned bottles, which were immediately acidified to pH 1.5 prior to analysis at laboratory. While maintaining oxide production under 2% (measured by CeO/Ce), samples were analyzed with ICP-MS tuned for maximum sensitivity using the high performance sample introduction system, which has the best signal to noise ratio. Correction of the raw instrument data was applied for variations in the internal standard, external drift, and interfering oxides. Hotplate digests for the US Geological Survey (USGS) dolerite reference material, W-2, were then utilized to calibrate the instrument response. A similar procedure was undertaken by Lawrence et al. [118] to assess the removal of paramagnetic Gd chelates through advanced water treatment plants. In the samples collected from the reverse osmosis (RO) permeate with an exceptionally clean matrix relative to most other sample types, LA concentrations were under the detection limit of the method. Thus, they considered evaporation (sub-boiling) as the most appropriate preconcentration method. In another investigation, a seaFAST 2 automated sample introduction system was interfaced with Q-ICP-MS to measure LA concentrations when approximately 25 ml of water samples was dried and the residue was treated with concentrated HNO_3, dried, and subsequently dissolved in 5% HNO_3 again [13].

Nevertheless, direct LA determination in water with ICP-MS is still challenging because of low LA concentrations, the matrix effect, and potential spectral interferences (isobaric and polyatomic). High salt contents in seawater result in signal suppression, instrument sensitivity drift, and clogging of the sample introduction and ions extraction systems. In addition, in plenty of environmental samples, some coexisting elements, such as Ba, Sn, and Sb, likely bring serious spectral interference (oxides ions and hydroxides) during the determination of La, Ce, and Eu by means of ICP-MS [120]. To overcome the problems and obtain more accurate and reliable analytical data through sensitive and interference-free determination, a preconcentration step regularly precedes the measurements.

Amongst various separation and preconcentration methods, approaches in which LAs are sorbed on different water-insoluble solid materials and further eluted with either acids or complexing reagents have been greatly used (Table 4). Solid phase extraction (SPE) procedures offer some important advantages in comparison with classical liquid—liquid extraction (LLE) techniques, including high enrichment factor, reduced organic solvents usage and exposure, rapid phase separation, and the feasibility of a combination with different detection techniques [121]. The other factors that make this technique very attractive for sample pretreatment are the various sorbent materials together with the broad range of chelating reagents and eluents. Furthermore, to appropriately perform sample pretreatment, they can be implemented and controlled in flow systems easily. Simple adsorption, ion-exchange, or complexation are the mechanisms of sorption, which are dependent on the nature of a given sorbent. Choosing solid material for the enrichment and removal of LAs ought to be based on the sample matrix, analyte concentration, and technique used for final detection whilst higher preconcentration can be achieved under adequate experimental conditions (time of sample loading, sorbent mass, volume of eluent) [120]. To accumulate Gd-diethylenetriaminepentaacetate (Gd-DTPA) on a stationary phase, the stability of the Gd chelates and their pH dependence are important. Efficiency of Chelex-100, Toyopearl AF Chelate-650 and C18 cartridges loaded with ethylhexylphosphates for the separation of LAs and Gd-DTPA from the water matrix was compared by Hennebrüder et al. [8]. Although the optimum acidity for LA sorption using Chelex 100 and Toyopearl was found to be at pH 5.5, Gd-DTPA retention is insufficient at this pH and depends on the complex concentration. However, both ionic LAs and Gd-DTPA are sufficiently retarded (about 70%) by C18 cartridges loaded with ethylhexylphosphates at pH$\sim$3 (the optimum acidity) [8]. Raju et al. [122] performed ICP-MS determination of Gd and its chelates after preconcentration with weak ion exchange material, Chelex-100; strong cation exchange material, Chromabond SA (SCX); and bis-(2-ethylhexyl)-phosphate (HDEHP) coated reverse phase C18. Of the three examined SPE materials, the HDEHP modified reverse phase C18 SPE was found to be the most efficient, yielding a hundredfold Gd preconcentration for contrast agents, such as Gd-BOPTA, Gd-DTPA, Gd-DOTA, and Gd-BT-DO3A [122]. Lawrence [123] used the LLE technique described by Lawrence and Kamber [124] (Table 4) to preconcentrate LAs in samples collected from the Brisbane River and Moreton Bay. Various preconcentration techniques employed for quantifying Gd_{anth} are compared with one another in Table 4.

Table 4. Different preconcentration procedures utilized for the separation of the complete set of LAs, including Gd_{anth}.

Preconcentration Technique	Solid Support	Complexing or Immobilizing Agent	Instrument	Analyte	Detection Limit (ng/L)	Preconcentration Factor	Time (min)	Samples	Reference
Solid phase extraction (SPE)	C18	Bis(2-ethylhexyl) Hydrogen Phosphate (HDEHP) and 2-Ethylhexyl Dihydrogen Phosphate (H2MEHP)	ICP-MS	LAs	-	200–1000	50–250 Batch method	Seawater	[125]
Solid phase extraction (SPE)	C18	Ethylhexylphosphates	ICP-MS	LAs	-	13.33	Batch method	River water	[8]
Solid phase extraction (SPE)	Chelex-100	-	ICP-MS	LAs	0.0012–0.034	100	120 Batch method	Seawater	[126]
Solid phase extraction (SPE)	Chelex-100	-	ICP-MS	LAs	50–560	20	Not given Batch method	River water	[8]
Solid phase extraction (SPE)	NOBIAS CHELATE PA1	-	HR-ICP-MS	LAs	0.01–0.82 pmol/kg	160–200	150 Batch method	Seawater	[127]
Countercurrent chromatography (CCC)	-	Di-2- ethylhexilphosporic acid (D2EHPA)	ICP-MS	LAs, Gd-DTPA	-	40	25	River water	[128]
Liquid-liquid extraction (LLE)	-	HDEHP (phosphoric acid 2-ethylhexyl ester -mono and di ester mixture)	ICP-MS	LAs	-	8	12	Seawater	[124]

Given that the concentrations of gadolinium diethylenetriaminepentaacetate (Gd-DTPA) and LAs in river water are mostly lower than the detection limits in quadrupole ICP-MS applying pneumatic nebulization, Hennebrüder et al. [128] developed a countercurrent chromatography (CCC) analytical scheme to enrich the analytes prior to ICP-MS detection (Table 4). The stationary phase was di-2-ethylhexylphosphoric acid (D2EHPA) dissolved in n-decane although the mobile phase was 0.01 M HNO_3. The employed CCC analytical results were comparable to those obtained by the preconcentration using SPE. The recovery rates for light and middle LAs (La to Tm) and Gd-DTPA (approximately 100% and 80%, respectively) were higher when the CCC scheme was used. Nonetheless, the SPE protocol yielded better recoveries for Yb and Lu (89% and 84%, respectively). To obtain the same preconcentration factor, the CCC technique required a shorter time frame (25 min for CCC vs. 120 min for SPE). The substantial SPE advantage that outweighs the disadvantage is its capability to preconcentrate several samples simultaneously [128].

To date, the method of Bau and Dulski [7], which is a slightly modified version of the procedure described by Shabani et al. [125] (Table 4), has been the most popular. Whereas the anthropogenic Gd, which is usually determined from the size of the Gd anomaly, is "bulk anthropogenic Gd", it is actually comprised of various chemical compounds used as MRI contrast agents. Contrary to Gd-DOPA (marketed over the last two decades) with a high complex stability, Gd-DTPA (the first on the market) is quantitatively sorbed to the phosphate ester used in the preconcentration procedure. Therefore, nowadays, the excess anthropogenic Gd determined following the original protocol of Bau and Dulski [7] underestimates the anomalous anthropogenic Gd in aqueous samples [129].

de Campos and Enzweiler [17] obtained the Gd concentration of filtered water samples with and without preconcentration. For the preconcentration of LAs, the procedure of Bau and Dulski [7] was used and the average time interval of 4–5 days was considered between the sampling (plus filtering and acidifying) and preconcentration. Interestingly, on average, the result obtained after preconcentration was 40% lower than that determined without this step. In another experiment, they preconcentrated LAs on the same day of sampling and the Gd content was determined by direct measurement in the filtrate, in the preconcentrated sample, and also in the chromatographic effluent (the sample percolated the SPE cartridge). Whilst the preconcentrated sample contained as low as 5% of the Gd concentration determined by direct analysis of the water sample, the chromatographic effluent sample accounted for 95% of it. The low Gd recovery in the eluate was attributed to incomplete dissociation of the paramagnetic Gd chelates in the short period of time between the pH adjustment of the filtered sample and performing the preconcentration. The unsatisfactory recovery of the dissolved anthropogenic Gd obtained indicates that the original protocol of Bau and Dulski [7] deserves further investigation and other reported anomalous Gd values [9,10,24,130] based on the same method might be underestimated [129].

To preconcentrate low LA concentrations of mussel shells and improve the decomposition of different contrast agents, Merschel and Bau [129] adjusted the method of Bau and Dulski [7] by adding approximately 20 mL of 30% H_2O_2 per 1 L of sample prior to passing the sample solution through an ion exchange column.

4.3. Speciation of Gd-CAs

It is vital to monitor Gd species rather than the total Gd concentrations to further investigate the anomalous Gd_{anth} and obtain more information about the processes and mechanisms resulting in Gd retention in organisms [97]. However, it is extremely challenging and likely impossible due to their low concentrations. Hence, a preconcentration step is occasionally inevitable.

A range of speciation techniques for the determination of different Gd-CAs in biological and environmental samples are compared in Table 5. With its simple quantitation approaches, ICP-MS is a sensitive, multielemental technique for the investigation of the overall metal content of various matrices. Nevertheless, all species data will be lost during atomization in the ICP. An effective strategy to regain such information is the hyphenation of separation techniques with ICP [131]. Although reversed phase chromatography (RP-HPLC) accounts for only very little retention for polar compounds, Mazzucotelli et al. [132] separated one contrast agent and several degradation products using RP-HPLC/ICP-MS. Moreover, size exclusion chromatography (SEC) was used with ICP-MS to separate free Gd^{3+} from Gd complexes and Gd adducts; however, the size differences between the individual Gd chelates are too small for effective SEC separation [133,134]. Ion chromatography (IC) [131,135], time-resolved luminescence (TRL) detector [136], and direct fluorescence [137] are also utilized for the separation of Gd species. Künnemeyer et al. [138] indicated a high separation efficiency of hydrophilic interaction chromatography (HILIC) for ionic and polar compounds by the determination of contrast agents in blood plasma using HILIC hyphenated to electrospray ionization mass spectrometry (ESI-MS). Since then, HILIC was applied in several scientific studies to analyze biological and environmental samples using zwitterionic, diol, and unboned silica-based stationary phases. Although ESI-MS and ICP-OES (inductively coupled plasma-optical emission spectroscopy) are other potential detection systems [138–141], the hyphenation of HILIC with ICP-MS has proven to be a powerful technique with the best performance for the determination of Gd-CAs, irrespective of aqueous solutions with quite high organic contents in HILIC [142–148]. Lindner et al. [148] employed HILIC-ICP-MS in an optimized speciation method and directly determined Gd-BT-DO3A, Gd-DOTA, and Gd- BOPTA in Berlin tap water, which have been previously found in surface waters of Berlin. By utilizing a sector field-based ICP-MS in the low-resolution mode and employing an ultra-sonic nebulization, Birka et al. [149] accomplished further improvements in sensitivities and decreased the detection limits for single species down to 0.9 ng/L, allowing the performance of speciation analysis in drinking water (Table 5). The analytical results of anthropogenic Gd speciation in aqueous samples are listed in Table 6.

Table 5. Speciation analysis for the determination of different Gd-CAs in biological and environmental samples.

Sample	Gd/Gd-CAs	Preconcentration/Separation	Instrument	LOD	Reference
Urine and hair	Gd, Gd-DTPA	SE separation: 20 mmol/L (pH = 7.4) of Tris-HCl buffer solutions was applied as eluents for separation of Gd^{3+} and Gd-DTPA standard solutions.	SEC-ICP-MS	3500 ng/L Gd	[134]
River water (Germany)	Gd, Gd–DTPA	C18-cartridges were loaded with ethylhexylphosphates and used for Gd preconcentration.	ICP-MS	200 ng/L Gd	[8]
Water (A bench-scale test)	Gd, Gd-DTPA, Gd-BT-DO3A	The chelates were separated using a Metrosep A Supp 3 - 250/4.6 anion separation column, while analysis of the displaced Gd^{3+} was carried out on a Nucleosil 5 SA-125/4.0 cation separation column using a 2-hydroxyisobutyric acid eluent.	IC-ICP-MS	-	[135]
Blood plasma	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	Separation of Gd chelates was carried out using a ZIC-HILIC column. For separations, eluent A of the mobile phase consisted of a solution of 12.5 mM ammonium formate and 12.5 mM formic acid in a mixture of 76/24 purified water/acetonitrile (pH of 3.75). Eluent B consisted of a solution of 12.5 mM ammonium formate and 12.5 mM formic acid in a mixture of 76/24 acetonitrile/water.	HILIC-ESI-MS	100 to 1000 nmol/L Gd	[138]
Blood serum	Gd-DTPA, Gd-BOPTA, Gd-DOTA, Gd-HP-DO3A, Gd-DTPA-BMA, Gd-DTPA-BMEA	Chromatographic separation was performed using a column suitable for the retention of hydrophilic polar compounds with isocratic elution. The mobile phase solution consisted of 10 mM ammonium acetate with 0, 0.5, 1, 2, 3 and 5% acetonitrile (pH 7.0 to 7.4).	HPLC-ICP-OES	8000 to 35,000 ng/L Gd	[140]
Hospital effluent, WWTP samples collected at different purification steps (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	The Gd contrast agents were separated using a zwitterionic ZIC-HILIC column. The mobile phase consisted of a solution of 12.5 mM ammonium formate and 12.5 mM formic acid in 76/24 acetonitrile/purified water (pH of 3.75).	HILIC-ICP-MS	1.0 nmol/L Gd	[142]
River water, lake water, WWTP influent and effluent (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	By the surface evaporation technique, water samples were concentrated to approximately 1:10 of the total volume without boiling. The Gd contrast agents were separated using a zwitterionic HILIC (ZIC-HILIC) column. The optimized mobile phase consisted of 20 mmol/L ammonium acetate in 60/40 acetonitrile/water (pH 5.8).	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]

Table 5. *Cont.*

Sample	Gd/Gd-CAs	Preconcentration/Separation	Instrument	LOD	Reference
Influent and effluent of a WWTP (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	ZIC-HILIC stationary phase was used for analysis of MRI contrast agents. In isocratic mode, the mobile phase consisted of a solution of 50 mM aqueous ammonium formate and acetonitrile (pH of 3.75)	HILIC-ICP-SFMS	130 ng/L Gd	[144]
Surface water (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DOTA, Gd-BT-DO3A	Separation of the Gd complexes was carried out in isocratic mode using an Accucore HILIC column with solid core particles. The mobile phase consisted of 50 mmol/L aqueous ammonium formate set to pH 3.75 with formic acid and acetonitrile.	HILIC-ICP-SFMS	80 to 100 pmol/L Gd	[145]
Teltow channel (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	To preconcentrate the contrast agents, samples were evaporated with IR light by soft heating to a fraction of about 1:20. The contrast agents were separated in isocratic mode using a ZIC-HILIC column. The mobile phase consisted of 20 mmol/L ammonium acetate in 65/35 acetonitrile/water (pH 7.3).	HILIC-ICP-MS	51 ± 11 ng/L Gd	[146]
Tap water (Germany)	Gd-DTPA, Gd-BOPTA, Gd-DTPA-BMA, Gd-DOTA, Gd-BT-DO3A	In an isocratic mode, the Gd complexes were separated using a zwitterionic ZIC-cHILIC column. The mobile phase consisted of 10 mmol/L ammonium acetate in 69/31 acetonitrile/water (pH 5.0).	HILIC-ICP-MS	1.4 to 3.5 ng/L Gd	[148]
Samples from waterworks (Germany)	Gd-DTPA, Gd-DOTA, Gd-BT-DO3A	An YMC-Triart diol HILIC column was used while the mobile phase consisted of 50 mmol/L aqueous ammonium formate (pH 3.7) and acetonitrile in isocratic mode.	HILIC-ICP-MS	8 to 14 pmol/L Gd	[149]

SEC-ICP-MS = Size-exclusion chromatography-inductively coupled plasma-mass spectrometry; IC-ICP-MS = Ion chromatography-inductively coupled plasma-mass spectrometry; HILIC-ESI-MS = Hydrophilic interaction chromatography-electrospray ionization-mass spectrometry; HPLC-ICP-OES = Hydrophilic interaction chromatography- inductively coupled plasma-optical emission spectroscopy; HILIC-ICP-MS = Hydrophilic interaction chromatography-inductively coupled plasma-mass spectrometry; HILIC-ICP-SFMS = Hydrophilic interaction chromatography-inductively coupled plasma sector field mass spectrometry

Table 6. Concentration of various Gd-based contrast agents in environmental samples. Total Gd was determined by means of ICP-MS.

Sample	Filtration	Gd-DTPA	Gd-BT-DO3A	Gd-BOPTA	Gd-DOTA	ΣGd Species	Total Gd	Instrument	LOD	Reference
Hospital effluent, Germany	0.45 µm	-	<D.L. [1]−23.1 nmol/L	-	-	<D.L.−23.1 nmol/L	0.5–27.9 nmol/L	HILIC-ICP-MS	1.0 nmol/L Gd	[142]
WWTP samples collected at different purification steps, Germany	0.45 µm	-	2.0–7.4 nmol/L	1.1–1.2 nmol/L	1.8–2.1 nmol/L	2.0–10.6 nmol/L	1.8–13.6 nmol/L	HILIC-ICP-MS	1.0 nmol/L Gd	[142]
Spree River (Dämeritzsee), Germany	0.45 µm	-	12.0 ng/L Gd	-	-	12.0 ng/L Gd	15.0 ng/L Gd	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]
Lake Wannsee, Germany	0.45 µm	106.0 ng/L Gd	184.0 ng/L Gd	12.0 ng/L Gd	110.0 ng/L Gd	412.0 ng/L Gd	436.0 ng/L Gd	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]
Havel River (Jungfernsee), Germany	0.45 µm	54.0 ng/L Gd	89.0 ng/L Gd	-	50.0 ng/L Gd	194.0 ng/L Gd	215.0 ng/L Gd	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]
WWTP influent, Germany	0.45 µm	29.0 ng/L Gd	36.0 ng/L Gd	-	34.0 ng/L Gd	99.0 ng/L Gd	122.0 ng/L Gd	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]
WWTP effluent, Germany	0.45 µm	27.0 ng/L Gd	34.0 ng/L Gd	-	36.0 ng/L Gd	97.0 ng/L Gd	118.0 ng/L Gd	HILIC-ICP-MS	22 ± 5 ng/L Gd	[143]
Surface water, Germany	0.2 µm	<D.L.−2.1 nmol/L	<D.L.−3.6 nmol/L	-	<D.L.−0.8 nmol/L	<D.L.−6.4 nmol/L	0.6–7.4 nmol/L	HILIC-ICP-SFMS	80–100 pmol/L Gd	[145]
WWTP effluent, Germany	0.2 µm	1.6 nmol/L	2.7 nmol/L	-	0.6 nmol/L	4.8 nmol/L	6.5 nmol/L	HILIC-ICP-SFMS	80–100 pmol/L Gd	[145]
Teltow channel, Germany	0.2 µm	-	18.0–471.0 ng/L Gd	-	16.0–456.0 ng/L Gd	34.0–926.0 ng/L Gd	50.0–990.0 ng/L Gd	HILIC-ICP-MS	51 ± 11 ng/L Gd	[146]
Tap water, Germany	0.2 µm	-	<D.L.−17.6 ng/L Gd	<D.L.−11.7 ng/L Gd	<D.L.−19.2 ng/L Gd	<D.L.−48.5 ng/L Gd	2.0–57.0 ng/L Gd	HILIC-ICP-MS	1.4–3.5 ng/L Gd	[148]
Surface water, Germany	0.2 µm	88.0–161.0 pmol/L	<D.L.−49.0 pmol/L	-	<D.L.−85.0 pmol/L	88.0–262.0 pmol/L	82.0–307.0 pmol/L	HILIC-ICP-MS	8–14 pmol/L Gd	[149]
Drinking water, Germany	0.2 µm	82.0–159.0 pmol/L	<D.L.	-	<D.L.−74.0 pmol/L	82.0–196.0 pmol/L	100.0–298.0 pmol/L	HILIC-ICP-MS	8–14 pmol/L Gd	[149]

[1] below detection limit.

4.4. Fractionation of Gd_{anth}

Generally, the mobility and bioavailability of LAs are governed by several factors, such as temperature, pH, as well as the concentration and type of organic and inorganic ligands [32,63]. According to Wells and Wells [150], nitrates, chlorides, and sulfates of LAs are soluble; however, their hydroxides, carbonates, and phosphates are insoluble. In river and estuarine waters, it is documented that pH, colloids, and salinity control the LA fractionation [32]. Researchers have generally used 0.2—0.45 µm membrane filters to obtain the so-called "dissolved" fraction (<0.4 µm). Both Elderfield et al. [151] and Sholkovitz [152] indicated that river colloids that can pass through the conventional membranes account for a large proportion of the "dissolved" LA pool. Hence, keeping the filtration protocol in mind is vital for the comparison of results [11].

Acidification experiments can be carried out to assess Gd fractionation [7,153]. In these experiments, the concentration of the 'dissolved' LAs (comprised of truly dissolved and colloidal LAs) is determined from the samples filtered through a 0.2 µm membrane filter prior to acidification. However, the sum of the dissolved plus the acid—soluble particulate LAs is detected from the samples acidified to pH 2 before 0.2 µm filtration. The results obtained from acidification experiments suggest [7,153]: (1) A difference between the speciation of Gd_{anth} in contaminated waters and the speciation of geogenic LAs in non-contaminated waters; (2) the presence of the excess Gd in the 'dissolved' LA fraction (<0.2 µm) rather than the acid—soluble particulate fraction; and (3) negligible particle reactivity of the anomalous Gd in water resources.

Filtration experiments are helpful in deciphering Gd fractionation as well. In addition to the 0.2 µm filtration, Kulaksız and Bau [24] passed an unacidified sample through a Vivaflow 50 unit (10 kDa MWCO) before preconcentration. LA concentrations <0.2 µm were referred to as 'dissolved' (i.e., the sum of truly dissolved LAs and LA bound to colloids), while concentrations in the <10 kDa LA pool were considered as "truly dissolved". Thus, the difference between the truly dissolved and the dissolved LA values explains the colloid-bound LA concentration. The result of the filtration experiment [24] indicates that the Gd_{anth} is not particle-reactive and exclusively present in the truly dissolved LA pool (<10 kDa), which is in accordance with the results of the acidification experiment carried out by Bau and Dulski [7].

4.5. Gd_{anth} in the Influent and Effluent of Wastewater Treatment Plants

The temporal variations in Gd values in influents and effluents depend on the dynamics between the percentage of the inhabitants undergoing MRI exams, the amount of Gd-CAs administered daily, the time of application and excretion of the contrast agents, and transport, dilution, and retention times in WWTPs. Because the predominant MRI scans with intravenous injection of Gd-CAs are performed in working days (i.e., between Monday and Friday) [144,154,155], Gd contents in both influents and effluents are weekday dependent [1]. In regions with no hospitals and limited access to the health system, anthropogenic Gd can also be determined since there might be people taking the exam in neighboring cities, resulting in the release of the stable Gd complexes when they get back home [16]. Möller et al. [3] attributed the considerable difference between the Gd_{anth} flux of WWTP effluents and the daily administered quantity of paramagnetic Gd chelates in Prague's hospitals to the excretion of 10% to 20% of the total Gd-CAs administered per day somewhere outside Prague. The discharge of treated wastewater into nearby natural water bodies can also play a significant role in a positive Gd anomaly in the LA pattern of water resources situated kilometers away from the point source. More than 50 km downstream of a WWTP source, Gd_{anth} (a tracer of treated wastewater) was easily detected in the receiving waters [2]. In areas where there are no major hospitals running MRI scans, enhanced Gd_{anth} contents in surface water resources with no connection to sewage treatment plants indicate an imperfect sewer system (uncontrolled dispersion by gutters and leaking sewers) as well [3]. Measured Gd values in wastewaters with Gd_{anth} are provided in Table 7.

Table 7. Total Gd concentrations and Gd anomalies reported in wastewaters and polluted aqueous samples. The subscript, SN, denotes the normalization to the Gd content in the continental crust (regularly PAAS was used) and the superscript * denotes the natural background value.

Sample	Study area	Filtration	Acidification	Instrument	Detection Limit	Gd	Gd_{SN}/Gd^{*}_{SN}	Reference
WWTP influent	GAW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	2040.0 pmol/L	111.0 [1]	[13]
	WSW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	126.0 pmol/L	4.2 [1]	[13]
	TAW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	3724.0 pmol/L	297.0 [1]	[13]
	JNW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	815.0 pmol/L	66.1 [1]	[13]
	AYW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	508.0 pmol/L	34.6 [1]	[13]
	Prague, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	437.0 nmol/m^3	35.0 [1]	[3]
WWTP effluent	GAW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	321.0 pmol/L	81.0 [1]	[13]
	WSW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	272.0 pmol/L	56.7 [1]	[13]
	TAW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	1286.0 pmol/L	271.0 [1]	[13]
	JNW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	786.0 pmol/L	186.0 [1]	[13]
	AYW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	676.0 pmol/L	147.0 [1]	[13]
	GPW, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	760.0 pmol/L	198.0 [1]	[13]
	São Paulo State, Brazil	0.22 µm	pH = 1.8–2.0	ICP-MS	1.4 ng/kg	86.0 ng/kg	9.5 [2]	[17]
	Brisbane, Australia	0.22 µm	pH = 1.5	ICP-MS	480 fmol/L	119.6–1795.0 pmol/kg	9.0–99.0 [3]	[16]
	Bremen–Seehausen, Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	1673.8 pmol/L	169.7 [4]	[15]
	Tallahassee, USA	Total	pH < 2	ICP-MS	-	0.12 µg/L	49.6 [1]	[156]
	Denver, USA	0.45 µm	pH < 2	ICP-MS	-	0.14 µg/L	20.0 [1]	[156]
	Boulder, USA	0.45 µm	pH < 2	ICP-MS	-	0.068 µg/L	62.7 [1]	[156]
	Berlin, Germany	0.2 µm	pH = 2.0	ICP-MS	0.01–0.1 pg/mL	201.0–7480.0 pmol/kg	64–2014 [1]	[153]
	Prague, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	253.0–278.0 nmol/m^3	55.4–61.7 [3]	[3]
	Berlin, Germany	0.2 mm	pH = 2.0	ICP-MS	-	7087 pmol/kg	1681 [1]	[7]
River water	Neuse River, USA	0.45 µm	-	Q-ICP-MS	-	25.4–342.0 pg/g	1.1–4.0 [5]	[19]
	Han River, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	45.9-209.0 pmol/L	1.6–8.2 [1]	[13]
	Gyeungan stream, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	141.0 pmol/L	8.7 [1]	[13]
	Tan stream, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	322.0 pmol/L	3.6 [1]	[13]
	Jungnang stream, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	308.0 pmol/L	7.4 [1]	[13]
	Gulpo stream, South Korea	0.2 µm	pH = 2.0	Q-ICP-MS	-	573.0 pmol/L	34.8 [1]	[13]
	Anhumas Creek, Brazil	0.22 µm	pH = 1.8–2.0	ICP-MS	1.4 ng/kg	13.0–207.0 ng/kg	1.1–86.7 [2]	[17]
	Havel River, Germany	0.2 µm	pH = 1.8–2.0	ICP-MS	-	3136 pmol/L	644.0 [2]	[10]
	Rhine River, Germany	0.2 µm	pH = 1.8–2.0	ICP-MS	-	11.8–188.0 ng/kg	4.4–110.0 [2]	[9]

Table 7. *Cont.*

Sample	Study area	Filtration	Acidification	Instrument	Detection Limit	Gd	Gd_{SN}/Gd^*_{SN}	Reference
	Danube River, Austria	0.2 µm	pH = 1.8–2.0	ICP-MS	-	48.6 pmol/L	2.3 [2]	[10]
	River Thames, England	0.2 µm	pH = 1.8–2.0	ICP-MS	-	28 pmol/L	2.5 [2]	[10]
	Weser River, Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	115.7 pmol/L	5.4 [4]	[15]
	Ems River, Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	154.0 pmol/L	3.6 [4]	[15]
	Elbe River, Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	94.8 pmol/L	5.5 [4]	[15]
	Susquehanna River, USA	0.2 µm	pH = 2.1	ICP-MS	0.01–0.1 pg/mL	62.6 pmol/L	2.2 [1]	[18]
	Shonai River, Japan	0.45 µm	pH = 1	ICP-MS	0.027 ng/L	8.9 ng/L	2.1 [1]	[12]
	Tenpaku River, Japan	0.45 µm	pH = 1	ICP-MS	0.027 ng/L	14.0 ng/L	9.4 [1]	[12]
	Tone River, Japan	0.04 µm	pH < 1.5	ICP-MS	-	10.3–58.5 pmol/kg	1.6–2.9 [1]	[11]
	Tama River, Japan	0.04 µm	pH < 1.5	ICP-MS	-	25.0–155.0 pmol/kg	2.2–3.3 [1]	[11]
	Ara River, Japan	0.04 µm	pH < 1.5	ICP-MS	-	26.2–66.2 pmol/kg	2.7–6.2 [1]	[11]
	Rokytka creek, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	35.7 $nmol/m^3$	1.8 [3]	[3]
	Wupper River, Germany	0.2 µm	pH = 1.9–2.1	ICP-MS	-	207.0 pmol/kg	30.0 [1]	[7]
	Spree River, Germany	0.2 µm	pH = 1.9–2.1	ICP-MS	-	43.1 pmol/kg	12.4 [1]	[7]
	Havel River, Germany	0.2 µm	pH = 1.9–2.1	ICP-MS	-	675 pmol/L	126 [2]	[7]
Seawater	Bahia Coast, NE Brazil	0.2 µm	pH = 1.8	HR ICP-MS	-	4.5–12.0 pmol/kg	1.0-3.4 [3]	[1]
	San Francisco Bay, USA	0.45 µm	pH = 1.8	HR ICP-MS	0.03 pmol/kg	14.1–171.4 pmol/kg	1.6–3.9 [3]	[20]
	Weser Estuary, Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	45.1–150.6 pmol/L	3.8–7.4 [4]	[15]
	North Sea	0.45 µm	pH = 1.8–2.0	ICP-MS	-	11.2–14.1 pmol/L	1.6–2.1 [4]	[15]
	Jade Bay, North Germany	0.45 µm	pH = 1.8–2.0	ICP-MS	-	17.9 pmol/L	1.7 [4]	[15]
	Nagoya port, Japan	0.45 µm	pH = 1	ICP-MS	0.027 ng/L	2.2 ng/L	1.6 [1]	[12]
	Tokyo Bay, Japan	0.04 µm	pH < 1.5	ICP-MS	-	8.4–15.4 pmol/kg	1.5–2.1 [1]	[11]
Lake water	Lake Paranoá, Brazil	0.2 µm	pH = 1.8–2.0	ICP-MS	-	7.9–35.2 ng/kg	18.5–40.9 [5]	[157]
	Kyjsky pond, Czech Republic	0.2 mm	pH = 2.0	ICP-MS	-	27.6 $nmol/m^3$	2.8 [3]	[3]
Groundwater	Honokohau Harbor well, Hawaii	0.45 µm	pH < 2	HR-ICP-MS	0.5–6 pmol/kg	192.0 pmol/kg	5.1 [1]	[21]
	Berlin, Germany	0.2 µm	pH = 2.0	ICP-MS	0.01–0.1 pg/ml	110.0–755.0 pmol/kg	3.6–53.0 [1]	[153]
Tap water	Western Berlin, Germany	0.2 µm	pH = 1.8–2.0	ICP-MS	-	3.5–115.0 pmol/L	1.4–33.7 [2]	[10]
	London, England	-	-	ICP-MS	-	12.3 pmol/L	1.5 [2]	[10]
	Berlin-Steglitz, Germany	0.2 µm	pH = 1.9–2.1	ICP-MS	-	39.8 pmol/kg	12.0 [1]	[7]

[1] $Gd_{SN}/Gd^*_{SN} = Gd_{SN}/(0.33Sm_{SN}+0.67Tb_{SN})$ [2] Calculated using linear regression method [3] Calculated using the third order polynomial fit [4] $Gd_{SN}/Gd^*_{SN} = Gd_{SN}/10\hat{}(2logSm\text{-}logNd)$ [5] Calculated from $logGd^*_{SN} = (4logEu_{SN}\text{-}logNd_{SN})/3$.

To enhance the settling characteristics of dissolved chemicals and suspended colloids, flocculation with Fe^{3+} or Al^{3+} salts is conducted in municipal sewage treatment plants. Since the added trivalent metal ions compete with Gd^{3+} for the organic ligands, depending on the thermodynamic complex stability, the trivalent metal salts encourage transmetallation (metal exchange) of the MRI Gd chelates and consequently release toxic Gd^{3+} [142,158]. For the retention of organically bound Gd, Möller et al. [3] suggested passing water from filter-beds and then treating it by FeOOH precipitation as a more effective approach than the FeOOH precipitation alone.

During regular treatment operations at WWTPs, a major proportion of natural dissolved LAs and only a minor fraction of Gd from anthropogenic sources can be removed, and the Gd_{anth} anomaly will consequently be highlighted [144,159]. Even with secondary or tertiary treatments (involving organic degradation and disinfection) in WWTPs, Gd-CAs are not expected to be treated owing to the high solubility of the polar or ionic Gd-CAs and the high stability of their complexes. Complexation of anthropogenic Gd from the MRI diagnostic system with either Cl^- or SO_4^{2-} and pronounced positive Gd anomalies in WWTP samples were reported by Song et al. [13]. Given the water treatment methods in the studied WWTPs, i.e., conventional activated sludge process (CASP), cilium nutrient removal (CNR), micro-chip filter (MCF), high class treatment (HCT), modified ludzack ettinger (MLE), and phased Isolation ditch (PID), they are probably not efficient enough for the removal of the Gd_{anth} load [13]. Telgmann et al. [144] simulated the aeration tank of a sewage treatment plant by batch experiments and showed that only approximately 10% of the Gd (Gd-DTPA) was removed during activated sludge treatment in WWTPs. Contrary to coagulation and microfiltration, the removal of 99.85% of anthropogenic Gd was reported by the reverse osmosis membrane [118].

On a population basis, in 2017, Pedreira et al. [1] calculated a total daily discharge of 11.5 ± 4.3 g Gd/d per 100,000 people and 15.5 ± 3.3 g Gd/d per 100,000 people for two WWTPs in Brazil. However, WWTPs in highly populated regions of Berlin were estimated to discharge 5 g Gd/d per 100,000 people from 1997 to 2000 [153], and a WWTP in southeast Queensland had a discharge of 3.3 g Gd/d per 100,000 people [16]. According to the estimation of Lawrence and Bariel [2], WWTPs in less populated areas accommodating approximately 100,000 persons are expected to discharge orders of magnitude lower that accounts for emissions of about 0.9 g Gd/d per 100,000 people. Pedreira et al. [1] explained that although they collected samples when the highest Gd inputs were expected (resulting in maximum values and not true averages), this discrepancy between the Gd emissions calculated for the WWTPs in Brazil and other cities may not be irrelevant to the upward trend of the number of MRI exams over the past 10 years (Figure 4). For instance, between 2007 and 2016, the statistics of MRI exams per 1000 American and Australian inhabitants increased by 32% and 124%, respectively. In Germany, the number of MRI exams increased from 84 exams/1000 inhabitants in 2007 to 136 exams/1000 inhabitants in 2015 (Figure 4) [160]. The recent statistics feasibly suggest the need for updating the estimates of Knappe et al. [153] and Lawrence and Bariel [1,2].

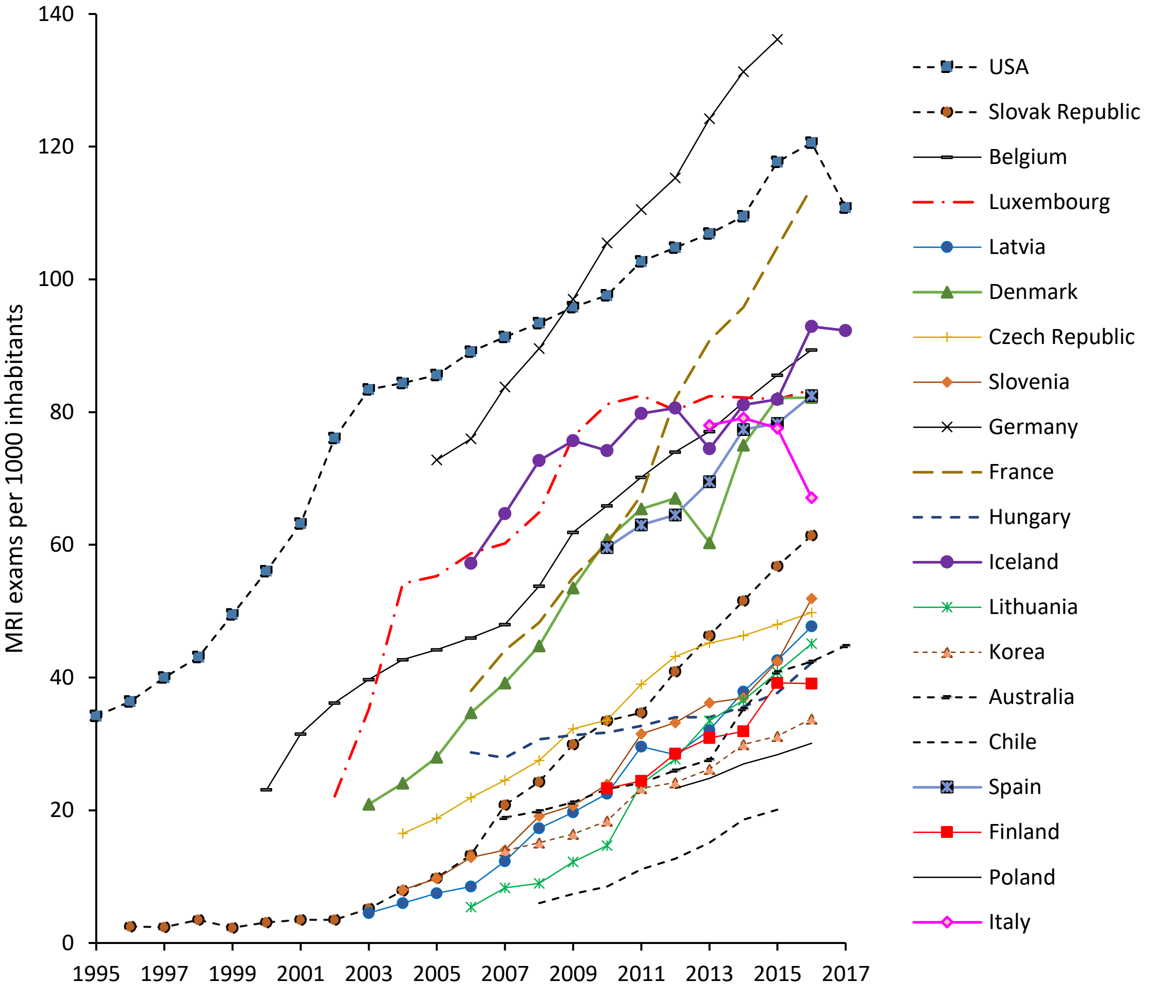


Figure 4. Long-term trend of the number of MRI exams per 1000 inhabitants performed in a range of countries [160].

4.6. Gd$_{anth}$ in Surface Water

Since the early report of anthropogenic Gd anomalies in river waters, which was traced back to the increasing use of Gd-CAs in the 1990s [7], similar positive Gd anomalies have been reported in the dissolved load of many urban rivers worldwide (Table 7). However, because anthropogenic Gd is less particle-reactive and more conservative, the contaminant is expected to be negligibly removed by colloidal organic matter and suspended particles during the estuarine mixing of river water and seawater compared to the removal of other elements in the LA series [7,11,161]. This further amplifies positive Gd anomaly of the riverine LA input into seawater and negatively affects the LA signature in coastal waters and in sea basins with restricted water circulation [7]. The difference in magnitude observed between rivers and marine water systems is likely related to the hydrodynamics that efficiently transports contaminants and dilutes the Gd$_{anth}$ signal in seawater (Table 7) [1].

Kulaksız and Bau [15] considered the pattern of water circulation and LA content in surface water and attributed anomalous Gd$_{anth}$ in the surface water of the southwestern North Sea off the coast of the East Frisian Islands to the Rhine River, the Ems River, and, feasibly, the Thames River. Besides rivers, ocean submarine outfalls provide a significant load of anthropogenic Gd in seawater. Notwithstanding the fact that plenty of coastal metropolises have adopted ocean submarine outfalls to discharge effluents into oceans after no treatment or only primary treatment, there has been no data regarding the impact of these pipeline systems on the concentrations of LAs in the marine environment until recently. Pedreira et al. [1] investigated the influence of submarine sewage outfalls on the distribution of LAs in the Atlantic coastal waters of northeast Brazil and observed the highest positive Gd anomalies in the vicinity of the pipeline system, which decreased with an increasing distance from the point source.

Documented Gd values and calculated positive Gd anomalies in contaminated water resources are listed in Table 7. In San Francisco Bay, increasing from approximately 10 pmol/kg in 1993 to 20 pmol/kg in 2000, anthropogenic Gd concentrations surged noticeably over the next 13 years and hit well above 100 pmol/kg in 2013. Surprisingly, between 2001 and 2013, the proportion of anthropogenic Gd was greater than 50% (Figure 5) [20].

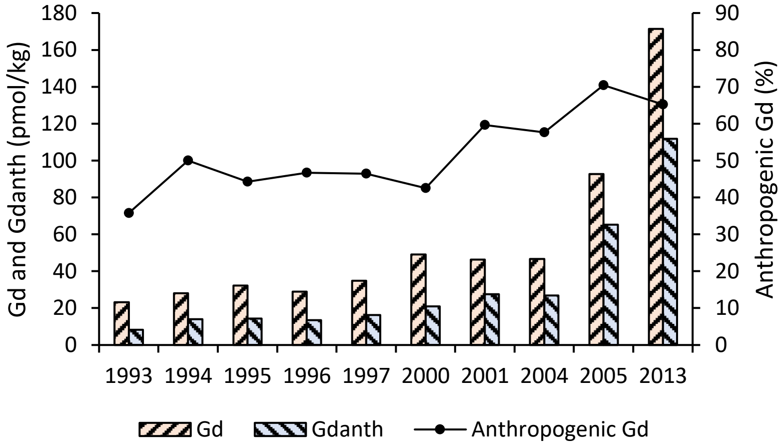


Figure 5. Gd concentration at the Lower South Bay, San Francisco Bay for the period from 1993 to 2013. Water samples were filtered through 0.45-µm membranes and acidified to pH = 1.8 [20].

When the LA concentrations are normalized, the proportion of LLAs (La to Eu), HLAs (Tb to Lu), and Gd (located between LLAs and HLAs) can be plotted on ternary diagrams to gain a better insight into water contamination by Gd-rich wastewater. Since anomalous Gd has been frequently reported in river waters, in this review, the diagrams are generated using the available analytical data for river waters collected at different parts of the world (Figure 6). The natural rivers included in Table 2 have no more than 10% Gd and their HLA concentrations are higher than those of LLAs, which is in accordance with Zhu et al. [12] (Figure 6a). In stark contrast, the proportion of Gd is greater than one-third in the majority of effluents and influents of WWTPs with substantial low LLAs (Figure 6b). Depending on the contamination degree and the wastewater percentage in river water, LA patterns can deviate from regular ones and various Gd proportions and HLAs/LLAs ratios can be observed in polluted river waters (Figure 6c).

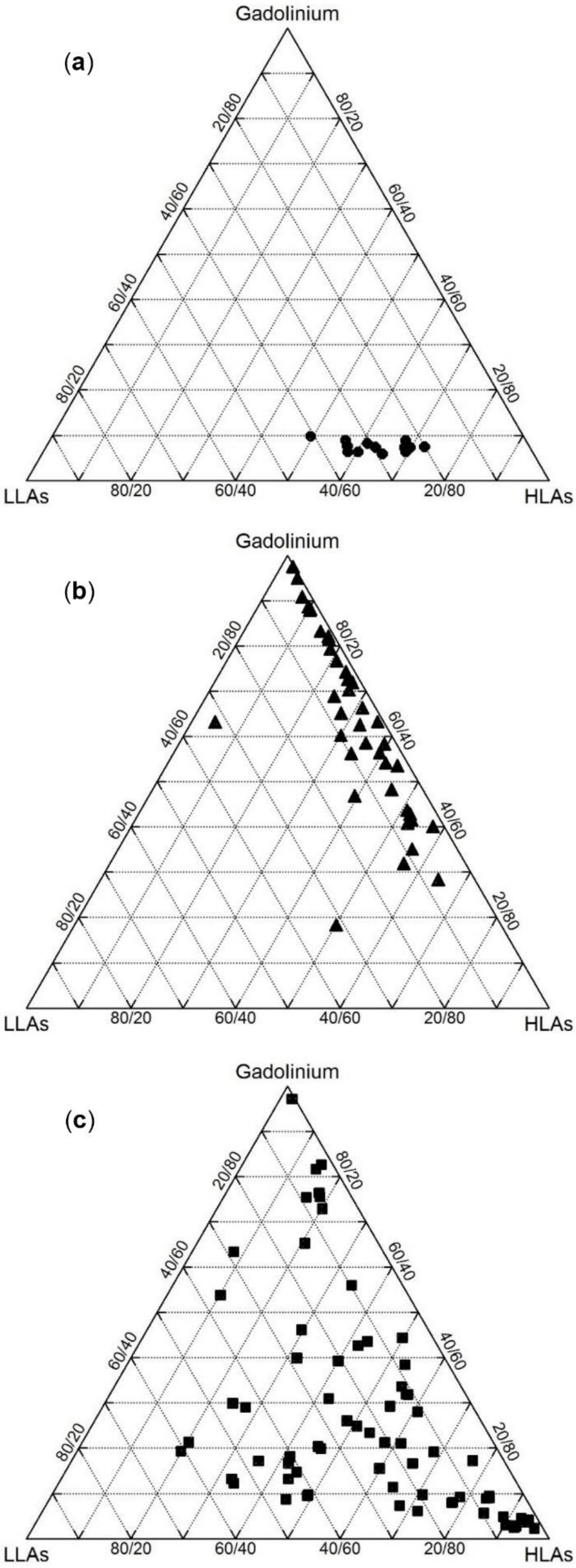


Figure 6. Comparison of LA concentrations (element contents are normalised to MUQ (Mud of Queensland; [87]) in: (**a**) Uncontaminated river waters; (**b**) influents and effluents of WWTPs; (**c**) contaminated river waters. The datasets of the articles listed in Tables 2 and 6 were used. LLAs: La to Eu, HLAs: Tb to Lu

4.7. Gd_{anth} and Surface Water—Groundwater Interaction

Without efficient sewage treatment processes and tap water purification stages, remarkable anthropogenic Gd can be discharged into surface water resources, transported to groundwater through natural and induced bank filtration, and subsequently transferred to tap water [10]. Gd complexes will not absorb, coprecipitate, or undergo ion exchange processes with organic and inorganic particulate matters, which are corresponded with the characteristics being crucial for their use as MRI diagnostic agents [3]. Positive Gd anomaly in LA distribution patterns of groundwater [21,153,162,163] show that processes, such as bank filtration (natural or induced), do not prevent the migration of Gd_{anth} into underground aquifers [10] (Table 7). Recently, attempts have been made to apply the Gd_{anth} content in groundwaters, river waters, and seawaters to hydrological studies [10,21,155,164].

Gd_{anth}, which is a refractory component, can be used as a conservative tracer of sewage effluent [5,15]. The changes of Gd_{anth} concentrations due to the mixing of bank filtrates and groundwater are well correlated with those of δD and $\delta^{18}O$, which shows the suitability of using Gd_{anth} for hydrological investigations at least on hydrological timescales of days to months [3]. Studying the source of groundwater by stable Gd complexes with a high residence time in the environment was also proposed by Möller et al. [162]. Bau and Dulski [7] evaluated the concentration of dissolved LAs (<0.2 μm fraction) in sewage treatment plant effluent, river water, and tap water in Berlin and indicated that the interaction of groundwater and surface water led to the pronounced positive Gd anomaly in tap water. By means of large positive Gd anomalies, treated wastewater from a local WWTP was traced from nearshore well waters into the coastal ocean, implying submarine groundwater discharge as the major source of LAs to the coastal water along the Kona Coast of the Big Island of Hawaii [21]. Further, over the last 15 years, Gd_{anth} content has increased more than four times in the Havel River in Berlin [10]. Considering the migration rate of water from the river to groundwater wells, which takes years to decades, Kulaksız and Bau [10] predicted an increase in nongeogenic Gd in western Berlin tap water over the next few years.

4.8. Human Health Risks of Gd_{anth}

Gd-CAs have been considered safe for a long period of time. However, potential Gd toxicity has been a health concern due to the Gd^{3+} ionic radius is roughly equal to Ca^{2+}, which enables the ion to block the calcium ion channels of cells [1]. In addition, Gd^{3+} regularly binds with a higher affinity than Ca^{2+} and competes with Ca^{2+} in physiological processes [165], resulting in a decrease of neuromuscular transmission, and interference in both intracellular enzymes and cell membrane processes through the transmetalation process [166,167]. Human beings would not survive 0.1 mmol/kg free Gd (e.g., $GdCl_2$) being injected into circulation [93].

Severe health concerns recently arose when it was discovered that Gd accumulation in human tissues could be followed by kidney failure [168], nephrogenic systemic fibrosis (NSF) [169,170], and anaphylactic shock [171], and results in death eventually [172]. Patients administered with Gd-CAs may have incorporated part of the Gd chelates into the brain, bone, and kidneys, and even trace amounts of Gd can trigger potentially fatal diseases, including NSF [173–177]. The common signs and symptoms of NSF are painful swelling, hardening and fibrosis under the epidermis with alterations of the normal pattern of collagen bundles as well as proliferation of spindle-shaped dermal cells [175,178]. The median time frame between NSF diagnosis and death was just below four months with a median age of 55 years [179]. The administered contrast agent excretion half-life is about 1.5 h in patients with functional kidneys, but may exceed this by a factor of 20 in those with renal insufficiency [180]. Given the anomalously high Gd concentrations in femoral heads of patients exposed to Gd-CAs (Omniscan and Prohance) 8 years prior to total hip replacement surgery, Darrah et al. [174] suggested the incorporation and retention of MRI Gd chelates into the bone for at least 8 years. The calculated Gd anomalies ranged up to more than 800 times the natural level.

Since the early reports of brain Gd deposition, many pediatric radiology departments have avoided using less stable linear Gd-CAs (Table 3) [111,112]. Although administrations of gadobutrol experienced a negligible rise, the statistics of other Gd-CAs dropped noticeably. The greatest decrease was reported in MRI tests using gadopentetate dimeglumine, which declined from 81% in 2010 to 15% in 2016. Gadoversetamide and gadodiamide were the least popular in 2010 (2% and 10%, respectively) and they were not used six years later anymore. The second most common Gd-CA was gadoterate meglumine in 2016. A considerable decrease of the number of NSF cases was observed after establishing a specific safety policy for the use of Gd-CAs (Table 3) by many scientific societies along with prominent health authorities, such as the Food and Drug Administration (FDA) and the European Medicines Agency (EMA). Furthermore, no new cases were reported after 2009, with the exception of rare isolated cases [181].

It is reported that the renal dysfunction lengthens the residence time of the paramagnetic Gd chelates inside the organism during which Gd-CAs probably undergo metabolic reactions, such as transmetallation and binding to endogenous compounds. Nonetheless, high level of Gd retention and NSF-like symptoms, such as hardening and swelling of the skin along with pain in different parts of the body, were experienced by some patients with normal renal function who underwent several MRI examinations [182–187]. The hypothesis is that Gd precipitates along with calcium phosphate and forms persistent insoluble deposits with diameters of a few micrometers, triggering the circulation of fibrocytes, which subsequently initiate fibrosis [188]. With their distinct advantages and limitations, a broad range of techniques, such as scanning electron microscopy/energy dispersive X-ray spectroscopy (SEM/EDS), transmission electron microscopy (TEM), secondary ion mass spectrometry (SIMS), extended X-ray absorption fine structure (EXAFS) spectroscopy, synchrotron X-ray fluorescence (SXRF) microscopy, and laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS), are available to perform elemental bio-imaging (EBI) and study insoluble Gd-containing deposits [175,189–193].

McDonald et al. [194] linked the hyperintense dentate nucleus (DN) on unenhanced MRI scans in patients reported by Kasahara et al. [195] to Gd accumulation in the brain area. Previously, Xia et al. [196] assumed that a disrupted blood-brain barrier (BBB) leads to Gd plaques in brain tumors originating from Gd-CAs. Nonetheless, brain Gd accumulation is not limited to those with either a damaged BBB or renal dysfunction [97,197]. Kanda et al. [198,199] found an association between high signal intensity in the brain of patients and frequent Gd-enhanced MRI applications in the past, irrespective of renal dysfunction. Brain Gd retention is not only persistent, but also observable in several parts of the brain, which has turned into a great health concern, resulting in the safety announcement of the Food and Drug Administration (FDA) [176].

4.9. Ecological Health Risks of Gd_{anth}

Because of high Gd_{anth} concentrations in water and Gd^{3+} toxicity, the potential bioacumulation of Gd chelates by both organisms and plants, which might lead to human exposure through the food chain, and factors (such as the concentration, speciation, and exposure duration) influencing the transport and distribution of Gd_{anth} in biota are worth investigating.

4.9.1. Organisms

There is some scientific evidence to support the uptake of natural and/or anthropogenic Gd by organisms. Despite the ubiquitous anthropogenic Gd anomalies in the Rhine River and the Weser River (Germany), the flat shale-normalized LA patterns of the aragonitic shell of a freshwater bivalve indicated that no anthropogenic Gd was incorporated into the shells and the difference in the speciation of geogenic and anthropogenic Gd in river waters was highlighted [129]. Furthermore, bioaccumulation of gadopentetic acid (Magnevist®) by a specific water flea (that are attractive for fish) exposed to polluted water and Gd containing nutrition algae (a marked nutrition for mussels) was recently studied using LA-ICP-MS as a promising tool for bio-imaging [200]. Whilst the exposure through the cultivation medium is followed by Gd signals on the skin and in the intestine, the uptake via nutrition algae led to the highest Gd intensities in the intestine. Therefore, there is a certain risk that the Gd-CA can enter the food chain of higher organisms and human beings as well [200]. Interestingly, various Gd sensitivities were documented across four sea urchins species (*Paracentrotus lividus*, *Arbacia lixula*, *Heliocidaris tuberculate*, and *Centrostephanus rodgersii*) by exposing their embryos to different concentrations of Gadolinium Acetate Tetrahydrate (GAT, Waco) [201]. Since gadolinium severely affected development of the embryos and skeleton growth of the sea urchins, the importance of Gd toxicity tests on several species for risk assessment was highlighted [201].

Recent studies have demonstrated Gd accumulation in the brain regions, liver, kidney, spleen, skin, and bone of different animals [197,202–205]. Bussi et al. [205] attributed various Gd accumulation rates in tissues to different Gd-CA washout rates. Biological retention of Gd is greatly dependent on Gd species. Non-ionic linear Gd-CAs pose the highest risk for potential Gd retention [147,182,197,206,207]. Contrary to the macrocyclic Gd-CA, gadoterate meglumine, the association of repeated administrations of linear Gd-CAs (gadobenate dimeglumine, gadopentetate dimeglumine, and gadodiamide) with brain Gd retention was reported in healthy rats [208]. According to Kartamihardja et al. [197], the tendency of Gd retention varies depending on the agent, regardless of renal function. Although renal failure increased short-term Gd retention after Gd-DTPA-BMA administration, it did almost not affect long-term Gd retention for Gd-CAs, implying that the chemical structures of accumulated Gd may be inconsistent and some Gd is retained initially prior to being removed slowly.

To determine the potential Gd^{3+} fetal toxicity after maternal gadoteridol intravenous injection, Oh et al. [209] obtained concentrations of the Gd chelate in nonhuman primate placenta, fetal tissues, and amniotic fluid. Compared to the maternal injected dose, the Gd complex concentrations in the fetal tissues and amniotic fluid were minimal, which may alleviate some concerns regarding MRI Gd chelate administration during pregnancy.

4.9.2. Plants

The mechanism of toxicity of lanthanides in plants depends on the La speciation, plant species, and growth stage. LAs may interrupt the uptake of various essential plant nutrients, in particular Ca with an almost similar ionic radii. As a result, Ca functions, including the formation of cell walls, root growth, photosynthesis, and flowering, might be interfered with [25]. In wheat, greater Gd^{3+} bioaccumulation was demonstrated in roots than in aerial parts [210].

Braun et al. [211] investigated four species of aquatic macrophytes (*Lemna gibba*, *Ceratophyllum demersum*, *Elodea nuttallii*, and *E. canadensis*) as potential biofilters for the removal of different Gd-CAs (Omniscan®and Dotarem®) from water. No significant Gd values in the tissues of the macrophytes revealed that the risk of Gd_{anth} accumulation in the food chain was low [211]. On the other hand, after the observation of a clear Gd signal in plants sampled downstream of a WWTP, Linder et al. [146] carried out research on the uptake of the Gd-containing contrast agents, Dotarem®, Gadovist®, and Multihance®, by the root system of cress plants under more controlled conditions. In a relatively short time frame, diffusion-driven transport or uptake of considerable amounts of Dotarem®and Gadovist®revealed an inefficiency of the casparian band (the filter system of the root, which normally prevents large molecules being taken up) for the Gd complexes. This might be due to the incomplete establishment of the casparian band in the young tips of the roots [212], leading to the uptake of the complexes through the apoplastic pathway [213]. Using speciation analyses, it was found that the contrast agents (not their metabolites) were taken up by the cress plant and transported to the leaves with no or only negligible modifications [146]. Moreover, uptake of Magnevist®and Omniscan®by plants cultivated in contaminated water or on contaminated soil irrigated with tap water was studied and accumulation of the Gd-CAs was demonstrated [200]. These findings contradict each other and indicate the need for further investigation to collect more information about the biological uptake and bioaccumulation of Gd-CAs by different plant species at different growth stages, and consequently the biomagnification of the microcontaminant through food chains.

5. Knowledge Gaps and Future Research Directions

Considering the aspects discussed in the previous sections, Figure 7 depicts different likely pathways through which MRI contrast agents pose a threat to human health. The research directions grouped into the following categories seem beneficial to bridge the knowledge gaps and ensure future ecological and human health.

5.1. Interventions to Prevent/Mitigate Environmental Pollution

- Little is known about the chemical behavior (degradation and transformation products) of Gd-CAs during the water treatment processes. A wide range of potential ecological and human health risks can be avoided by allocating more financial resources to investigate and upgrade current inefficient wastewater treatment technologies and water purification techniques.

5.2. Detecting Gd_{anth} as a Microcontaminant in Aqueous Samples and Collecting a Reliable Dataset

- With the ongoing controversy over securing LA resources and future affordable LA supplies, the environmental aspects are no longer a priority for the industrialized nations. Currently, no extensive LA dataset exists to serve as a background level for monitoring studies, which might be due to a lack of knowledge about their human health risks and ecotoxicology [130]. The dataset should be completed with temporal and spatial data about the distribution of a range of Gd species in the hydrologic cycle in order to empower decision makers to protect the environment.
- Considering the diversity of preconcentration procedures and the lack of a certified scheme for the recently marketed Gd-CAs, proposing a harmonized method to determine anthropogenic Gd seems necessary. A method accepted by a large group of scientists can facilitate the comparison of data reported from different parts of the world.
- Since various MRI Gd chelates account for the Gd_{anth}, the capability of the existing preconcentration procedures to efficiently extract Gd complexes that will be marketed in the future should be verified to prevent underestimation of the anomalous anthropogenic Gd in aqueous samples.

5.3. Hydrological Studies and Monitoring Other Microcontaminants in Water Resources

- Those living in small cities may undergo MRI tests in medical centers of neighboring metropolises and excrete the injected Gd-CAs when they get back home. It shows the possibility of investigating either water pollution or surface water—groundwater interactions in cities without MRI centers.
- When the anomalous Gd content can be traced back to WWTPs, it is expected that other emerging microcontaminants, including pharmaceuticals and personal care products (PPCPs), that cannot be completely removed during water treatment also occur in water [214]. Using statistical and machine learning techniques, it would be possible to identify a relationship between different microcontaminants and predict the presence of other chemical constituents without expensive chemical analysis.

5.4. Ecological and Human Health Risks of Gd_{anth}

- Given the complexity of aquatic environments comprised of different biological systems (e.g., plants, algae, zooplankton, and fish) that differently interact with Gd chelates, there is a risk that Gd-CAs can enter the food chain of higher organisms and human beings. Thus, Gd bioavailability, long-term bioaccumulation of Gd-CAs in the biosphere, their stability in the environment or in biological systems, and the toxicological impacts of Gd complexes are urgent issues that need to be addressed. The uptake of Gd chelates and their bioaccumulation in different parts of plants cultivated in farms and irrigated with polluted water deserve more attention as well. The abovementioned studies might also help to find hyper-accumulator species for Gd_{anth}.

- Environmentally relevant concentrations of Gd-containing MRI contrast agents should be considered for further ecotoxicological investigations. In addition, the mixture effect of anthropogenic Gd and other organic and inorganic stressors on ecotoxicity warrants further research. Thus, the results can be reliably extrapolated to complex environmental systems where Gd_{anth} simultaneously exists with numerous chemical compounds.
- In comparison with speciation analysis, the determination of total LA contents to calculate anomalous positive Gd has been more common. Since the stability of Gd-CAs varies substantially and biological Gd retention is remarkably dependent on Gd species, it makes sense to consider various Gd-CAs to interpret their fate in the environment and assess their health risk comprehensively.
- The conventional methodologies (e.g., ICP-OES and ICP-MS) that have been often hyphenated to separation systems, such as HPLC, to quantify the total Gd content or Gd species in different biological tissues provide no spatial resolved information and therefore the precise Gd bioaccumulation remains unknown. The application of recent analytical techniques (including LA-ICP-MS) to create images of the elemental distribution at the cellular range in organisms and plants (bio-imaging) will be advantageous.
- Prevailing winds and ocean currents may transport Gd-containing WWTP effluents towards shores during the high season [1,215]. Given that river water and seawater might be used for drinking purposes or recreational activities, local residents (especially fishermen) and tourists are exposed to anthropogenic Gd through dermal and ingestion pathways. Hence, evaluations of the health risk through chronic and acute exposure scenarios for children and adults would be beneficial.
- Irrespective of renal insufficiency or damaged BBB, Gd retention was observed in the human body after the administration of paramagnetic Gd chelates. Therefore, the investigation of Gd deposition in the human body and the corresponding health problems ought to be stressed.

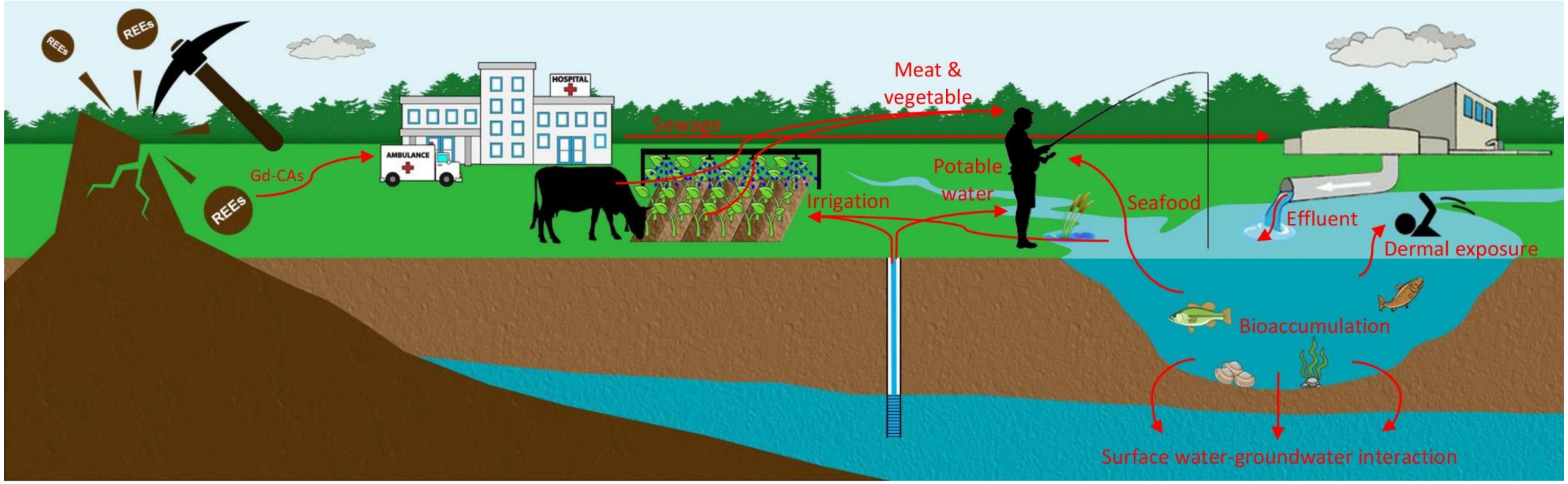


Figure 7. A simplified diagram demonstrating various probable pathways that excreted gadolinium chelates can enter our food chain. To date, anthropogenic Gd has been detected in hospital sewage, WWTP effluent, surface water, groundwater, and tap water. Bioaccumulation of various Gd-CAs in different aquatic species should be further investigated to reach unanimous agreement. Whilst the accumulation of Gd contrast agent has not been reported in fish, vegetables, and livestock, it is worth investigating to enhance our well-being. Besides, exposure through dermal and ingestion pathways should be evaluated.

6. Conclusions

Many publications and research projects were presented in this review. It was indicated that stable MRI contrast agents have been negligibly removed from wastewater, and discharged into surface water, resulting in the detection of Gd_{anth} in river water, seawater, groundwater, and tap water. Moreover, the number of MRI exams is on the increase worldwide, leading to greater Gd_{anth} concentrations in water resources in the future. Part of the administered Gd complexes have been incorporated in the brain, bone, and kidneys of patients, and even trace amounts of them can encourage potentially fatal health problems, such as NSF. Although paramagnetic Gd chelates have been detected as a microcontaminant in water resources since the mid-1990s, their uptake and bioaccumulation by different biota have not been thoroughly investigated and, as a result, no permissible limit has been introduced by authorities to screen safe food and water. It is probable that every individual takes up minute daily amounts of Gd via the consumption of contaminated drinking water and ingredients (meat, seafood, and vegetables) that is considerably higher than the natural doses. Consequently, it is crucial to monitor Gd species in the environment regularly and promote life-long observations to assess their ecotoxicity and adverse human health effects. Cooperation between medical scientists and environmental experts is highly important to achieve further progress. The discrimination against bioaccumulation of nongeogenic Gd by at least some organisms and plants corroborates the previous evidence revealing long environmental half-lifes of the paramagnetic Gd chelates. Moreover, it adds to the growing body of evidence suggesting conservative anthropogenic Gd as a useful tracer of WWTP effluents in river water, lake water, seawater, groundwater and tap water. Evaluating the distribution of MRI Gd-CAs also offers a rather inexpensive yet robust way of monitoring xenobiotics, such as pharmaceuticals and personal care products (PPCPs), in water. Although many highly interesting and relevant questions are still to be addressed, the presented analytical methods constitute a perfect base for future research.

Author Contributions: Visualization, P.E.; writing—original draft preparation, P.E.; writing—review and editing, M.B.

Funding: This research received no external funding.

Acknowledgments: The authors would like to express their gratitude towards the anonymous reviewers for careful reading of the manuscript and the constructive comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pedreira, R.M.; Pahnke, K.; Böning, P.; Hatje, V. Tracking hospital effluent-derived gadolinium in Atlantic coastal waters off Brazil. *Water Res.* **2018**, *145*, 62–72. [CrossRef] [PubMed]
2. Lawrence, M.G.; Bariel, D.G. Tracing treated wastewater in an inland catchment using anthropogenic gadolinium. *Chemosphere* **2010**, *80*, 794–799. [CrossRef] [PubMed]
3. Möller, P.; Paces, T.; Dulski, P.; Morteani, G. Anthropogenic Gd in surface water, drainage system, and the water supply of the city of Prague, Czech Republic. *Environ. Sci. Technol.* **2002**, *36*, 2387–2394. [CrossRef] [PubMed]
4. Fuganti, A.; Möller, P.; Morteani, G.; Dulski, P. Gadolinio ed altre terre rare usabili come traccianti per stabilire l'eta il movimento ed i rischi delle acque sotterranee: Esempio dell'area di Trento. *Geol. Tec. Ambien* **1996**, *4*, 13–18.
5. Möller, P.; Morteani, G.; Dulski, P. Anomalous gadolinium, cerium, and yttrium contents in the adige and isarco river waters and in the water of their tributaries (Provinces Trento and Bolzano/Bozen, NE Italy). *Acta Hydrochim. Hydrobiol.* **2003**, *31*, 225–239. [CrossRef]
6. Elbaz-Poulichet, F.; Seidel, J.L.; Othoniel, C. Occurrence of an anthropogenic gadolinium anomaly in river and coastal waters of Southern France. *Water Res.* **2002**, *36*, 1102–1105. [CrossRef]
7. Bau, M.; Dulski, P. Anthropogenic origin of positive gadolinium anomalies in river waters. *Earth Planet. Sci. Lett.* **1996**, *143*, 245–255. [CrossRef]

8. Hennebrüder, K.; Wennrich, R.; Mattusch, J.; Stärk, H.-J.; Engewald, W. Determination of gadolinium in river water by SPE preconcentration and ICP-MS. *Talanta* **2004**, *63*, 309–316. [CrossRef] [PubMed]
9. Kulaksız, S.; Bau, M. Rare earth elements in the Rhine River, Germany: First case of anthropogenic lanthanum as a dissolved microcontaminant in the hydrosphere. *Environ. Int.* **2011**, *37*, 973–979. [CrossRef] [PubMed]
10. Kulaksız, S.; Bau, M. Anthropogenic gadolinium as a microcontaminant in tap water used as drinking water in urban areas and megacities. *Appl. Geochem.* **2011**, *26*, 1877–1885. [CrossRef]
11. Nozaki, Y.; Lerche, D.; Alibo, D.S.; Tsutsumi, M. Dissolved indium and rare earth elements in three Japanese rivers and Tokyo Bay: Evidence for anthropogenic Gd and In. *Geochim. Cosmochim. Acta* **2000**, *64*, 3975–3982. [CrossRef]
12. Zhu, Y.; Hoshino, M.; Yamada, H.; Itoh, A.; Haraguchi, H. Gadolinium anomaly in the distributions of rare earth elements observed for coastal seawater and river waters around Nagoya City. *Bull. Chem. Soc. Jpn.* **2004**, *77*, 1835–1842. [CrossRef]
13. Song, H.; Shin, W.-J.; Ryu, J.-S.; Shin, H.S.; Chung, H.; Lee, K.-S. Anthropogenic rare earth elements and their spatial distributions in the Han River, South Korea. *Chemosphere* **2017**, *172*, 155–165. [CrossRef] [PubMed]
14. Nozaki, Y.; Lerche, D.; Alibo, D.S.; Snidvongs, A. The estuarine geochemistry of rare earth elements and indium in the Chao Phraya River, Thailand. *Geochim. Cosmochim. Acta* **2000**, *64*, 3983–3994. [CrossRef]
15. Kulaksız, S.; Bau, M. Contrasting behaviour of anthropogenic gadolinium and natural rare earth elements in estuaries and the gadolinium input into the North Sea. *Earth Planet. Sci. Lett.* **2007**, *260*, 361–371. [CrossRef]
16. Lawrence, M.G.; Ort, C.; Keller, J. Detection of anthropogenic gadolinium in treated wastewater in South East Queensland, Australia. *Water Res.* **2009**, *43*, 3534–3540. [CrossRef]
17. de Campos, F.F.; Enzweiler, J. Anthropogenic gadolinium anomalies and rare earth elements in the water of Atibaia River and Anhumas Creek, Southeast Brazil. *Environ. Monit. Assess.* **2016**, *188*, 281. [CrossRef]
18. Bau, M.; Knappe, A.; Dulski, P. Anthropogenic gadolinium as a micropollutant in river waters in Pennsylvania and in Lake Erie, northeastern United States. *Chemie der erde-Geochemistry* **2006**, *66*, 143–152. [CrossRef]
19. Smith, C.; Liu, X.-M. Spatial and temporal distribution of rare earth elements in the Neuse River, North Carolina. *Chem. Geol.* **2018**, *488*, 34–43. [CrossRef]
20. Hatje, V.; Bruland, K.W.; Flegal, A.R. Increases in anthropogenic gadolinium anomalies and rare earth element concentrations in San Francisco Bay over a 20 year record. *Environ. Sci. Technol.* **2016**, *50*, 4159–4168. [CrossRef]
21. Johannesson, K.H.; Palmore, C.D.; Fackrell, J.; Prouty, N.G.; Swarzenski, P.W.; Chevis, D.A.; Telfeyan, K.; White, C.D.; Burdige, D.J. Rare earth element behavior during groundwater–seawater mixing along the Kona Coast of Hawaii. *Geochim. Cosmochim. Acta* **2017**, *198*, 229–258. [CrossRef]
22. Schwarzenbach, R.P.; Escher, B.I.; Fenner, K.; Hofstetter, T.B.; Johnson, C.A.; Von Gunten, U.; Wehrli, B. The challenge of micropollutants in aquatic systems. *Science* **2006**, *313*, 1072–1077. [CrossRef] [PubMed]
23. Ternes, T.; Joss, A. *Human Pharmaceuticals, Hormones and Fragrances*; IWA Publishing: London, UK, 2007.
24. Kulaksız, S.; Bau, M. Anthropogenic dissolved and colloid/nanoparticle-bound samarium, lanthanum and gadolinium in the Rhine River and the impending destruction of the natural rare earth element distribution in rivers. *Earth Planet. Sci. Lett.* **2013**, *362*, 43–50. [CrossRef]
25. Gwenzi, W.; Mangori, L.; Danha, C.; Chaukura, N.; Dunjana, N.; Sanganyado, E. Sources, behaviour, and environmental and human health risks of high-technology rare earth elements as emerging contaminants. *Sci. Total Environ.* **2018**, *636*, 299–313. [CrossRef] [PubMed]
26. Merian, E.; Anke, M.; Ihnat, M.; Stoeppler, M. *Elements and Their Compounds in the Environment: Occurrence, Analysis and Biological Relevance*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2004.
27. Kabata-Pendias, A. *Trace Elements in Soils and Plants*; CRC press: Florida, FL, USA, 2010.
28. Kabata-Pendias, A.; Mukherjee, A.B. *Trace Elements from Soil to Human*; Springer Science & Business Media: Berlin, Germany, 2007.
29. Sonich-Mullin, C. *Rare Earth Elements: A Review of Production, Processing, Recycling, and Associated Environmental Issues*; EPA/600/R-12/572; Office of Research and Development: Washington, DC, USA, 2013; 135p.
30. Migaszewski, Z.M.; Gałuszka, A. The characteristics, occurrence, and geochemical behavior of rare earth elements in the environment: A review. *Crit. Rev. Environ. Sci. Technol.* **2015**, *45*, 429–471. [CrossRef]

31. Ramos, S.J.; Dinali, G.S.; Oliveira, C.; Martins, G.C.; Moreira, C.G.; Siqueira, J.O.; Guilherme, L.R. Rare earth elements in the soil environment. *Curr. Pollut. Rep.* **2016**, *2*, 28–50. [CrossRef]
32. Sholkovitz, E.R. The aquatic chemistry of rare earth elements in rivers and estuaries. *Aquat. Geochem.* **1995**, *1*, 1–34. [CrossRef]
33. Turetta, C.; Barbaro, E.; Capodaglio, G.; Barbante, C. Dissolved rare earth elements in the central-western sector of the Ross Sea, Southern Ocean: Geochemical tracing of seawater masses. *Chemosphere* **2017**, *183*, 444–453. [CrossRef]
34. Petrosino, P.; Sadeghi, M.; Albanese, S.; Andersson, M.; Lima, A.; De Vivo, B. REE contents in solid sample media and stream water from different geological contexts: Comparison between Italy and Sweden. *J. Geochem. Explor.* **2013**, *133*, 176–201. [CrossRef]
35. Li, X.-F.; Chen, Z.-B.; Chen, Z.-Q. Distribution and fractionation of rare earth elements in soil–water system and human blood and hair from a mining area in southwest Fujian Province, China. *Environ. Earth Sci.* **2014**, *72*, 3599–3608. [CrossRef]
36. Byrne, R.H.; Li, B. Comparative complexation behavior of the rare earths. *Geochim. Cosmochim. Acta* **1995**, *59*, 4575–4589. [CrossRef]
37. Enghag, P. *Encyclopedia of the Elements: Technical Data-History-Processing-Applications*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008.
38. U.S. Geological Survey. *Metal Prices in the United States through 2010: U.S. Geological Survey Scientific Investigations Report 2012-5188*; U.S. Geological Survey: Reston, Virginia, USA, 2013; p. 204.
39. Hedrick, J.B. Rare earths. In *Minerals Year Book—2000*; U. S. Geological Survey: Washington, DC, USA, 2000; pp. 62.1–62.10.
40. Alonso, E.; Sherman, A.M.; Wallington, T.J.; Everson, M.P.; Field, F.R.; Roth, R.; Kirchain, R.E. Evaluating rare earth element availability: A case with revolutionary demand from clean technologies. *Environ. Sci. Technol.* **2012**, *46*, 3406–3414. [CrossRef] [PubMed]
41. Du, X.; Graedel, T. Uncovering the global life cycles of the rare earth elements. *Sci. Rep.* **2011**, *1*, 145. [CrossRef] [PubMed]
42. Massari, S.; Ruberti, M. Rare earth elements as critical raw materials: Focus on international markets and future strategies. *Resourc. Policy* **2013**, *38*, 36–43. [CrossRef]
43. U.S. Geological Survey. *Rare Earth Statistics, in Kelly, T.D., and Matos, G.R., comps., Historical Statistics for Mineral and Material Commodities in the United States: U.S. Geological Survey Data Series 140*; US Geological Survey: Reston, Virginia, USA, 2014. Available online: http://minerals.usgs.gov/minerals/pubs/historical-statistics/ (accessed on 7 December 2018).
44. Ahonen, S.; Arvanitidis, N.; Auer, A.; Baillet, E.; Bellato, N.; Binnemans, K.; Blengini, G.A.; Bonato, D.; Brouwer, E.; Brower, S. *Strengthening The European Rare Earths Supply-Chain Challenges And Policy Options a Report by the European Rare Earths Competency Network (Erecon)*; Technical Report; European Commission: Maastricht, The Netherlands, 2015.
45. Newton, D.E.; Edgar, K.J. *Chemical Elements*; UXL (A Part of Gale, Cengage Learning): Michigan, MI, USA, 2010.
46. Bahl, C.R.H.; Nielsen, K.K. The effect of demagnetization on the magnetocaloric properties of gadolinium. *J. Appl. Phys.* **2009**, *105*, 013916. [CrossRef]
47. Humphries, M. *Rare Earth Elements: The Global Supply Chain*; Congressional Research Service: Washington, DC, USA, 2013.
48. Krishnamurthy, N.; Gupta, C.K. *Extractive Metallurgy of Rare Earths*; CRC Press: Florida, FL, USA, 2015.
49. Aime, S.; Cabella, C.; Colombatto, S.; Geninatti Crich, S.; Gianolio, E.; Maggioni, F. Insights into the use of paramagnetic Gd (III) complexes in MR-molecular imaging investigations. *J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Resonan. Med.* **2002**, *16*, 394–406. [CrossRef] [PubMed]
50. Babar, S.; Saifuddin, A. MRI of the post-discectomy lumbar spine. *Clin. Radiol.* **2002**, *57*, 969–981. [CrossRef]
51. Bencardino, J.T.; Palmer, W.E. Imaging of hip disorders in athletes. *Radiol. Clin. N. Am.* **2002**, *40*, 267–287, vi-vii. [CrossRef]
52. Rajagopalan, S.; Prince, M. Magnetic resonance angiographic techniques for the diagnosis of arterial disease. *Cardiol. Clin.* **2002**, *20*, 501–512. [CrossRef]

53. Choudhury, L.; Mahrholdt, H.; Wagner, A.; Choi, K.M.; Elliott, M.D.; Klocke, F.J.; Bonow, R.O.; Judd, R.M.; Kim, R.J. Myocardial scarring in asymptomatic or mildly symptomatic patients with hypertrophic cardiomyopathy. *J. Am. College Cardiol.* **2002**, *40*, 2156–2164. [CrossRef]
54. Conaghan, P.G.; O'connor, P.; McGonagle, D.; Astin, P.; Wakefield, R.J.; Gibbon, W.W.; Quinn, M.; Karim, Z.; Green, M.J.; Proudman, S. Elucidation of the relationship between synovitis and bone damage: A randomized magnetic resonance imaging study of individual joints in patients with early rheumatoid arthritis. *Arthritis Rheum.* **2003**, *48*, 64–71. [CrossRef] [PubMed]
55. Degani, H.; Chetrit-Dadiani, M.; Bogin, L.; Furman-Haran, E. Magnetic resonance imaging of tumor vasculature. *Thromb. Haemost.* **2003**, *89*, 25–33. [PubMed]
56. Stwertka, A. *A Guide to the Elements*; Oxford University Press: New York, NY, USA, 2002.
57. Taylor, K.C. Automobile catalytic converters. In *Catalysis*; Springer: Berlin/Heidelberg, Germany, 1984; pp. 119–170.
58. Hurst, C. *China's Rare Earth Elements Industry: What Can the West Learn?* Institute for the Analysis of Global Security: Washington, DC, USA, 2010.
59. Reimann, C.; De Caritat, P. *Chemical Elements in the Environment: Factsheets for the Geochemist and Environmental Scientist*; Springer-Verlag: Berlin/Heidelberg, Germany, 1998.
60. Korzh, V.D. *Geochemistry of Elemental Composition of the Hydrosphere*; Nauka: Moscow, Russia, 1991. (In Russian)
61. Bibak, A.; Stümp, S.; Knudsen, L.; Gundersen, V. Concentrations of 63 elements in cabbage and sprouts in Denmark. *Commun. Soil Sci. Plant Anal.* **1999**, *30*, 2409–2418. [CrossRef]
62. Eriksson, J. *Concentrations of 61 Trace Elements in Sewage Sludge, Farmyard Manure, Mineral Fertiliser, Precipitation and in Oil and Crops*; Swedish Environmental Protection Agency: Stockholm, Sweden, 2001; Volume 5159.
63. Byrne, R.; Sholkovitz, E. Marine chemistry and geochemistry of the lanthanides. In *Handbook on the Physics and Chemistry of Rare Earths*; North-Holland: Amsterdam, Netherlands, 1996; Volume 23, pp. 497–593.
64. Yan, Z.; Liu, G.; Sun, R.; Tang, Q.; Wu, D.; Wu, B.; Zhou, C. Geochemistry of rare earth elements in groundwater from the Taiyuan Formation limestone aquifer in the Wolonghu Coal Mine, Anhui province, China. *J. Geochem. Explor.* **2013**, *135*, 54–62. [CrossRef]
65. Lawrence, M.G.; Greig, A.; Collerson, K.D.; Kamber, B.S. Direct quantification of rare earth element concentrations in natural waters by ICP-MS. *Appl. Geochem.* **2006**, *21*, 839–848. [CrossRef]
66. Braun, J.-J.; Pagel, M.; Muller, J.-P.; Bilong, P.; Michard, A.; Guillet, B. Cerium anomalies in lateritic profiles. *Geochim. Cosmochim. Acta* **1990**, *54*, 781–795. [CrossRef]
67. Sholkovitz, E.R.; Landing, W.M.; Lewis, B.L. Ocean particle chemistry: The fractionation of rare earth elements between suspended particles and seawater. *Geochim. Cosmochim. Acta* **1994**, *58*, 1567–1579. [CrossRef]
68. Leleyter, L.; Probst, J.-L.; Depetris, P.; Haida, S.; Mortatti, J.; Rouault, R.; Samuel, J. Distribution des terres rares dans les sédiments fluviaux: Fractionnement entre les phases labiles et résiduelles. In *Comptes rendus de l'Académie des sciences. Série 2. Sciences de la terre et des planètes*; 1999; Volume 329, pp. 45–52.
69. Cidu, R.; Biddau, R. Transport of trace elements under different seasonal conditions: Effects on the quality of river water in a Mediterranean area. *Appl. Geochem.* **2007**, *22*, 2777–2794. [CrossRef]
70. Aubert, D.; Stille, P.; Probst, A. REE fractionation during granite weathering and removal by waters and suspended loads: Sr and Nd isotopic evidence. *Geochim. Cosmochim. Acta* **2001**, *65*, 387–406. [CrossRef]
71. Leybourne, M.I.; Johannesson, K.H. Rare earth elements (REE) and yttrium in stream waters, stream sediments, and Fe–Mn oxyhydroxides: Fractionation, speciation, and controls over REE+ Y patterns in the surface environment. *Geochim. Cosmochim. Acta* **2008**, *72*, 5962–5983. [CrossRef]
72. Stallard, R.; Edmond, J. Geochemistry of the Amazon: 2. The influence of geology and weathering environment on the dissolved load. *J. Geophys. Res. Oceans* **1983**, *88*, 9671–9688. [CrossRef]
73. Meybeck, M. Global chemical weathering of surficial rocks estimated from river dissolved loads. *Am. J. Sci.* **1987**, *287*, 401–428. [CrossRef]
74. Tosiani, T.; Loubet, M.; Viers, J.; Valladon, M.; Tapia, J.; Marrero, S.; Yanes, C.; Ramirez, A.; Dupre, B. Major and trace elements in river-borne materials from the Cuyuni basin (southern Venezuela): Evidence for organo-colloidal control on the dissolved load and element redistribution between the suspended and dissolved load. *Chem. Geol.* **2004**, *211*, 305–334. [CrossRef]

75. Möller, P.; Knappe, A.; Dulski, P. Seasonal variations of rare earths and yttrium distribution in the lowland Havel River, Germany, by agricultural fertilization and effluents of sewage treatment plants. *Appl. Geochem.* **2014**, *41*, 62–72. [CrossRef]
76. Haley, B.A.; Frank, M.; Hathorne, E.; Pisias, N. Biogeochemical implications from dissolved rare earth element and Nd isotope distributions in the Gulf of Alaska. *Geochim. Cosmochim. Acta* **2014**, *126*, 455–474. [CrossRef]
77. Fisher, A.; Kara, D. Determination of rare earth elements in natural water samples—A review of sample separation, preconcentration and direct methodologies. *Anal. Chim. Acta* **2016**, *935*, 1–29. [CrossRef] [PubMed]
78. Gaillardet, J.; Viers, J.; Dupré, B. Trace elements in river waters. *Treatise Geochem.* **2003**, *5*, 605.
79. Alibo, D.S.; Nozaki, Y. Rare earth elements in seawater: Particle association, shale-normalization, and Ce oxidation. *Geochim. Cosmochim. Acta* **1999**, *63*, 363–372. [CrossRef]
80. Lawrence, M.G.; Kamber, B.S. The behaviour of the rare earth elements during estuarine mixing—Revisited. *Mar. Chem.* **2006**, *100*, 147–161. [CrossRef]
81. McLennan, S.M. Rare earth elements in sedimentary rocks; influence of provenance and sedimentary processes. *Rev. Mineral. Geochem.* **1989**, *21*, 169–200.
82. Taylor, S.R.; McClennan, S.M. *The Continental Crust: Its Composition and Evolution*; Blackwell: Boston, MA, USA, 1985.
83. Rudnick, R.L.; Gao, S. Composition of the continental crust. *Treatise Geochem.* **2003**, *3*, 659.
84. Sholkovitz, E.R. Rare earth elements in the sediments of the North Atlantic Ocean, Amazon Delta, and East China Sea; reinterpretation of terrigenous input patterns to the oceans. *Am. J. Sci.* **1988**, *288*, 236–281. [CrossRef]
85. Johannesson, K.H.; Stetzenbach, K.J.; Hodge, V.F.; Kreamer, D.K.; Zhou, X. Delineation of Ground-Water Flow Systems in the Southern Great Basin Using Aqueous Rare Earth Element Distributions. *Groundwater* **1997**, *35*, 807–819. [CrossRef]
86. Cidu, R.; Antisari, L.V.; Biddau, R.; Buscaroli, A.; Carbone, S.; Da Pelo, S.; Dinelli, E.; Vianello, G.; Zannoni, D. Dynamics of rare earth elements in water–soil systems: The case study of the Pineta San Vitale (Ravenna, Italy). *Geoderma* **2013**, *193*, 52–67. [CrossRef]
87. Kamber, B.S.; Greig, A.; Collerson, K.D. A new estimate for the composition of weathered young upper continental crust from alluvial sediments, Queensland, Australia. *Geochim. Cosmochim. Acta* **2005**, *69*, 1041–1058. [CrossRef]
88. Lawrence, M.G.; Jupiter, S.D.; Kamber, B.S. Aquatic geochemistry of the rare earth elements and yttrium in the Pioneer River catchment, Australia. *Mar. Freshw. Res.* **2006**, *57*, 725–736. [CrossRef]
89. Haraguchi, H.; Itoh, A.; Kimata, C.; Miwa, H. Speciation of yttrium and lanthanides in natural water by inductively coupled plasma mass spectrometry after preconcentration by ultrafiltration and with a chelating resin. *Analyst* **1998**, *123*, 773–778. [CrossRef]
90. Itoh, A.; Nagasawa, T.; Zhu, Y.; Lee, K.-H.; FUJIMORI, E.; HARAGUCHI, H. Distributions of major-to-ultratrace elements among the particulate and dissolved fractions in natural water as studied by ICP-AES and ICP-MS after sequential fractionation. *Anal. Sci.* **2004**, *20*, 29–36. [CrossRef] [PubMed]
91. Masuda, A.; Ikeuchi, Y. Lanthanide tetrad effect observed in marine environment. *Geochem. J.* **1979**, *13*, 19–22. [CrossRef]
92. Nozaki, Y.; Zhang, J. The rare earth elements and yttrium in the coastal/offshore mixing zone of Tokyo Bay waters and the Kuroshio. In *Biogeochemical Processes and Ocean Flux in the Western Pacific*; Terra Scientific Publishing Company: Tokyo, Japan, 1995; pp. 171–184.
93. Thomsen, H.S. Are the increasing amounts of gadolinium in surface and tap water dangerous? *Acta Radiol.* **2017**, *58*, 259–263. [CrossRef]
94. Holzbecher, E.; Knappe, A.; Pekdeger, A. Identification of degradation characteristics–exemplified by Gd–DTPA in a large experimental column. *Environ. Model. Assess.* **2005**, *10*, 1–8. [CrossRef]
95. USEPA. *Risk Assessment Guidance for Superfund. Volume I: Human Health Evaluation Manual (Part A)*; EPA/540/1-89/002; U.S. Environmental Protection Agency: Washington, DC, USA, 1989.
96. Kanal, E. Gadolinium based contrast agents (GBCA): Safety overview after 3 decades of clinical experience. *Magn. Reson. Imaging* **2016**, *34*, 1341–1345. [CrossRef] [PubMed]

97. Clases, D.; Sperling, M.; Karst, U. Analysis of metal-based contrast agents in medicine and the environment. *TrAC Trends Anal. Chem.* **2018**, *104*, 135–147. [CrossRef]
98. Weinmann, H.-J.; Brasch, R.C.; Press, W.-R.; Wesbey, G.E. Characteristics of gadolinium-DTPA complex: A potential NMR contrast agent. *Am. J. Roentgenol.* **1984**, *142*, 619–624. [CrossRef]
99. Alexander, V. Design and synthesis of macrocyclic ligands and their complexes of lanthanides and actinides. *Chem. Rev.* **1995**, *95*, 273–342. [CrossRef]
100. Runge, V.M. Commentary on T1-weighted hypersignal in the deep cerebellar nuclei after repeated administrations of gadolinium-based contrast agents in healthy rats: Difference between linear and macrocyclic agents. *Investig. Radiol.* **2015**, *50*, 481–482. [CrossRef] [PubMed]
101. Idée, J.M.; Port, M.; Raynal, I.; Schaefer, M.; Le Greneur, S.; Corot, C. Clinical and biological consequences of transmetallation induced by contrast agents for magnetic resonance imaging: A review. *Fundam. Clin. Pharmacol.* **2006**, *20*, 563–576. [CrossRef] [PubMed]
102. Port, M.; Idée, J.-M.; Medina, C.; Robic, C.; Sabatou, M.; Corot, C. Efficiency, thermodynamic and kinetic stability of marketed gadolinium chelates and their possible clinical consequences: A critical review. *Biometals* **2008**, *21*, 469–490. [CrossRef]
103. Fraum, T.J.; Ludwig, D.R.; Bashir, M.R.; Fowler, K.J. Gadolinium-based contrast agents: A comprehensive risk assessment. *J. Magn. Reson. Imaging* **2017**, *46*, 338–353. [CrossRef]
104. Hao, D.; Ai, T.; Goerner, F.; Hu, X.; Runge, V.M.; Tweedle, M. MRI contrast agents: Basic chemistry and safety. *J. Magn. Reson. Imaging* **2012**, *36*, 1060–1071. [CrossRef]
105. Uggeri, F.; Aime, S.; Anelli, P.L.; Botta, M.; Brocchetta, M.; de Haeen, C.; Ermondi, G.; Grandi, M.; Paoli, P. Novel contrast agents for magnetic resonance imaging. Synthesis and characterization of the ligand BOPTA and its Ln (III) complexes (Ln = Gd, La, Lu). X-ray structure of disodium (TPS-9-145337286-CS)-[4-carboxy-5, 8, 11-tris (carboxymethyl)-1-phenyl-2-oxa-5, 8, 11-triazatridecan-13-oato (5-)] gadolinate (2-) in a mixture with its enantiomer. *Inorg. Chem.* **1995**, *34*, 633–643.
106. Caravan, P.; Ellison, J.J.; McMurry, T.J.; Lauffer, R.B. Gadolinium (III) chelates as MRI contrast agents: Structure, dynamics, and applications. *Chem. Rev.* **1999**, *99*, 2293–2352. [CrossRef] [PubMed]
107. Brücher, E.; Sherry, A. Stability and toxicity of contrast agents. In *The Chemistry of Contrast Agents in Medical Magnetic Resonance Imaging*; John Wiley & Sons: West Sussex, UK, 2001; Volume 2.
108. Caravan, P.; Comuzzi, C.; Crooks, W.; McMurry, T.; Choppin, G.; Woulfe, S. Thermodynamic stability and kinetic inertness of MS-325, a new blood pool agent for magnetic resonance imaging. *Inorg. Chem.* **2001**, *40*, 2170–2176. [CrossRef] [PubMed]
109. Bellin, M.; Vasile, M.; Morel-Precetti, S. Currently used non-specific extracellular MR contrast media. *Eur. Radiol.* **2003**, *13*, 2688–2698. [CrossRef] [PubMed]
110. Moreau, J.; Guillon, E.; Pierrard, J.C.; Rimbault, J.; Port, M.; Aplincourt, M. Complexing Mechanism of the Lanthanide Cations Eu3+, Gd3+, and Tb3+ with 1, 4, 7, 10-Tetrakis (carboxymethyl)-1, 4, 7, 10-tetraazacyclododecane (dota)—Characterization of Three Successive Complexing Phases: Study of the Thermodynamic and Structural Properties of the Complexes by Potentiometry, Luminescence Spectroscopy, and EXAFS. *Chem. Eur. J.* **2004**, *10*, 5218–5232. [PubMed]
111. Malayeri, A.A.; Brooks, K.; Bryant, L.H.; Evers, R.; Kumar, P.; Reich, D.S.; Bluemke, D.A. NIH perspective on reports of gadolinium deposition in the brain. *J. A. College Radiol. JACR* **2016**, *13*, 237–241. [CrossRef] [PubMed]
112. American College of Radiology-American Society of Neuroradiology. ACR-ASNR Position Statement on the Use of Gadolinium Contrast Agents. Available online: https://www.asnr.org/wpcontent/uploads/2017/03/ACR_ASNR_Position_Statement_on_the_Use_of_Gadolinium_Contrast_Agents.pdf (accessed on 14 August 2017).
113. European Society of Urogenital Radiology. Available online: http://esur-cm.org/index.php/en/a-general-adverse-reactions-2 (accessed on 14 February 2019).
114. Yang, X.-C.; Sachs, F. Block of stretch-activated ion channels in Xenopus oocytes by gadolinium and calcium ions. *Science* **1989**, *243*, 1068–1071. [CrossRef] [PubMed]
115. Telgmann, L.; Sperling, M.; Karst, U. Determination of gadolinium-based MRI contrast agents in biological and environmental samples: A review. *Anal. Chim. Acta* **2013**, *764*, 1–16. [CrossRef] [PubMed]
116. Bau, M.; Dulski, P.; Moller, P. Yttrium and holmium in South Pacific seawater: Vertical distribution and possible fractionation mechanisms. *Oceanogr. Lit. Rev.* **1995**, *11*, 955.

117. Behrens, M.K.; Muratli, J.; Pradoux, C.; Wu, Y.; Böning, P.; Brumsack, H.-J.; Goldstein, S.L.; Haley, B.; Jeandel, C.; Paffrath, R. Rapid and precise analysis of rare earth elements in small volumes of seawater-Method and intercomparison. *Mar. Chem.* **2016**, *186*, 110–120. [CrossRef]
118. Lawrence, M.G.; Keller, J.; Poussade, Y. Removal of magnetic resonance imaging contrast agents through advanced water treatment plants. *Water Sci. Technol.* **2010**, *61*, 685–692. [CrossRef] [PubMed]
119. Zuccato, E.; Calamari, D.; Natangelo, M.; Fanelli, R. Presence of therapeutic drugs in the environment. *Lancet* **2000**, *355*, 1789–1790. [CrossRef]
120. Pyrzynska, K.; Kubiak, A.; Wysocka, I. Application of solid phase extraction procedures for rare earth elements determination in environmental samples. *Talanta* **2016**, *154*, 15–22. [CrossRef]
121. Augusto, F.; Hantao, L.W.; Mogollón, N.G.; Braga, S.C. New materials and trends in sorbents for solid-phase extraction. *TrAC Trends Anal. Chem.* **2013**, *43*, 14–23. [CrossRef]
122. Raju, C.S.K.; Lück, D.; Scharf, H.; Jakubowski, N.; Panne, U. A novel solid phase extraction method for pre-concentration of gadolinium and gadolinium based MRI contrast agents from the environment. *J. Anal. Atomic Spectrom.* **2010**, *25*, 1573–1580.
123. Lawrence, M.G. Detection of anthropogenic gadolinium in the Brisbane River plume in Moreton Bay, Queensland, Australia. *Mar. Pollut. Bull.* **2010**, *60*, 1113–1116. [CrossRef] [PubMed]
124. Lawrence, M.G.; Kamber, B.S. Rare earth element concentrations in the natural water reference materials (NRCC) NASS-5, CASS-4 and SLEW-3. *Geostand. Geoanal. Res.* **2007**, *31*, 95–103. [CrossRef]
125. Shabani, M.B.; Akagi, T.; Masuda, A. Preconcentration of trace rare-earth elements in seawater by complexation with bis (2-ethylhexyl) hydrogen phosphate and 2-ethylhexyl dihydrogen phosphate adsorbed on a C18 cartridge and determination by inductively coupled plasma mass spectrometry. *Anal. Chem.* **1992**, *64*, 737–743. [CrossRef]
126. Sawatari, H.; Toda, T.; Saizuka, T.; Kimata, C.; Itoh, A.; Haraguchi, H. Multielement determination of rare earth elements in coastal seawater by inductively coupled plasma mass spectrometry after preconcentration using chelating resin. *Bull. Chem. Soc. Jpn.* **1995**, *68*, 3065–3070. [CrossRef]
127. Hatje, V.; Bruland, K.W.; Flegal, A.R. Determination of rare earth elements after pre-concentration using NOBIAS-chelate PA-1®resin: Method development and application in the San Francisco Bay plume. *Mar. Chem.* **2014**, *160*, 34–41. [CrossRef]
128. Hennebrüder, K.; Engewald, W.; Stärk, H.-J.; Wennrich, R. Enrichment of rare-earth elements (REE) and Gd-DTPA in surface water samples by means of countercurrent chromatography (CCC). *Anal. Chim. Acta* **2005**, *542*, 216–221. [CrossRef]
129. Merschel, G.; Bau, M. Rare earth elements in the aragonitic shell of freshwater mussel Corbicula fluminea and the bioavailability of anthropogenic lanthanum, samarium and gadolinium in river water. *Sci. Total Environ.* **2015**, *533*, 91–101. [CrossRef] [PubMed]
130. Kulaksiz, S. Rare Earth Elements as Emerging Contaminants in the Rhine River, Germany and Its Tributaries. Ph.D. Thesis, Jacobs University Bremen, Bremen, Germany, 2012.
131. Meermann, B.; Sperling, M. Hyphenated techniques as tools for speciation analysis of metal-based pharmaceuticals: Developments and applications. *Anal. Bioanal. Chem.* **2012**, *403*, 1501–1522. [CrossRef] [PubMed]
132. Mazzucotelli, A.; Bavastello, V.; Magi, E.; Rivaro, P.; Tomba, C. Analysis of gadolinium polyaminopolycarboxylic complexes by HPLC–ultrasonic nebulizer–ICP-AES hyphenated technique. *Anal. Proc. incl. Anal. Commun.* **1995**, *32*, 165–167. [CrossRef]
133. Krüger, R.; Braun, K.; Pipkorn, R.; Lehmann, W.D. Characterization of a gadolinium-tagged modular contrast agent by element and molecular mass spectrometry. *J. Anal. Atomic Spectrom.* **2004**, *19*, 852–857. [CrossRef]
134. Loreti, V.; Bettmer, J. Determination of the MRI contrast agent Gd-DTPA by SEC–ICP–MS. *Anal. Bioanal. Chem.* **2004**, *379*, 1050–1054. [CrossRef] [PubMed]
135. Pfundstein, P.; Martin, C.; Schulz, W.; Ruth, K.M.; Wille, A.; Moritz, T.; Steinbach, A.; Flottmann, D. IC-ICP-MS analysis of gadolinium-based MRI contrast agents. *LC GC N. Am.* **2011**, *29*, 27–29.
136. Moutiez, E.; Prognon, P.; Mahuzier, G.; Bourrinet, P.; Zehaf, S.; Dencausse, A. Time-resolved Luminescence as a Novel Detection Mode for the Simultaneous High-performance Liquid Chromatographic Determination of Gadolinium–DOTA and Gd 3+. *Analyst* **1997**, *122*, 1347–1352. [CrossRef]
137. Hagan, J.J.; Taylor, S.C.; Tweedle, M.F. Fluorescence detection of gadolinium chelates separated by reversed-phase high-performance liquid chromatography. *Anal. Chem.* **1988**, *60*, 514–516. [CrossRef]

138. Künnemeyer, J.; Terborg, L.; Nowak, S.; Scheffer, A.; Telgmann, L.; Tokmak, F.; Günsel, A.; Wiesmüller, G.; Reichelt, S.; Karst, U. Speciation analysis of gadolinium-based MRI contrast agents in blood plasma by hydrophilic interaction chromatography/electrospray mass spectrometry. *Anal. Chem.* **2008**, *80*, 8163–8170. [CrossRef] [PubMed]
139. Künnemeyer, J.; Terborg, L.; Nowak, S.; Brauckmann, C.; Telgmann, L.; Albert, A.; Tokmak, F.; Krämer, B.K.; Günsel, A.; Wiesmüller, G.A. Quantification and excretion kinetics of a magnetic resonance imaging contrast agent by capillary electrophoresis-mass spectrometry. *Electrophoresis* **2009**, *30*, 1766–1773. [CrossRef] [PubMed]
140. Kahakachchi, C.L.; Moore, D.A. Speciation of gadolinium in gadolinium-based magnetic resonance imaging agents by high performance liquid chromatography inductively coupled plasma optical emission spectrometry. *J. Anal. Atomic Spectrom.* **2009**, *24*, 1389–1396. [CrossRef]
141. Kahakachchi, C.L.; Moore, D.A. Identification and characterization of gadolinium (III) complexes in biological tissue extracts. *Metallomics* **2010**, *2*, 490–497. [CrossRef]
142. Künnemeyer, J.; Terborg, L.; Meermann, B.r.; Brauckmann, C.; Möller, I.; Scheffer, A.; Karst, U. Speciation analysis of gadolinium chelates in hospital effluents and wastewater treatment plant sewage by a novel HILIC/ICP-MS method. *Environ. Sci. Technol.* **2009**, *43*, 2884–2890. [CrossRef]
143. Raju, C.S.K.; Cossmer, A.; Scharf, H.; Panne, U.; Lück, D. Speciation of gadolinium based MRI contrast agents in environmental water samples using hydrophilic interaction chromatography hyphenated with inductively coupled plasma mass spectrometry. *J. Anal. Atomic Spectrom.* **2010**, *25*, 55–61. [CrossRef]
144. Telgmann, L.; Wehe, C.A.; Birka, M.; Künnemeyer, J.; Nowak, S.; Sperling, M.; Karst, U. Speciation and isotope dilution analysis of gadolinium-based contrast agents in wastewater. *Environ. Sci. Technol.* **2012**, *46*, 11929–11936. [CrossRef] [PubMed]
145. Birka, M.; Wehe, C.A.; Telgmann, L.; Sperling, M.; Karst, U. Sensitive quantification of gadolinium-based magnetic resonance imaging contrast agents in surface waters using hydrophilic interaction liquid chromatography and inductively coupled plasma sector field mass spectrometry. *J. Chromatogr. A* **2013**, *1308*, 125–131. [CrossRef] [PubMed]
146. Lindner, U.; Lingott, J.; Richter, S.; Jakubowski, N.; Panne, U. Speciation of gadolinium in surface water samples and plants by hydrophilic interaction chromatography hyphenated with inductively coupled plasma mass spectrometry. *Anal. Bioanal. Chem.* **2013**, *405*, 1865–1873. [CrossRef] [PubMed]
147. Birka, M.; Wentker, K.S.; Lusmöller, E.; Arheilger, B.; Wehe, C.A.; Sperling, M.; Stadler, R.; Karst, U. Diagnosis of nephrogenic systemic fibrosis by means of elemental bioimaging and speciation analysis. *Anal. Chem.* **2015**, *87*, 3321–3328. [CrossRef] [PubMed]
148. Lindner, U.; Lingott, J.; Richter, S.; Jiang, W.; Jakubowski, N.; Panne, U. Analysis of gadolinium-based contrast agents in tap water with a new hydrophilic interaction chromatography (ZIC-cHILIC) hyphenated with inductively coupled plasma mass spectrometry. *Anal. Bioanal. Chem.* **2015**, *407*, 2415–2422. [CrossRef]
149. Birka, M.; Wehe, C.A.; Hachmöller, O.; Sperling, M.; Karst, U. Tracing gadolinium-based contrast agents from surface water to drinking water by means of speciation analysis. *J. Chromatogr. A* **2016**, *1440*, 105–111. [CrossRef] [PubMed]
150. Wells, W.H.; Wells, V.L. The Lanthanides, Rare Earth Elements. *Patty's Toxicol.* **2001**, 817–840. [CrossRef]
151. Elderfield, H.; Upstill-Goddard, R.; Sholkovitz, E. The rare earth elements in rivers, estuaries, and coastal seas and their significance to the composition of ocean waters. *Geochim. Cosmochim. Acta* **1990**, *54*, 971–991. [CrossRef]
152. Sholkovitz, E.R. Chemical evolution of rare earth elements: Fractionation between colloidal and solution phases of filtered river water. *Earth Planet. Sci. Lett.* **1992**, *114*, 77–84. [CrossRef]
153. Knappe, A.; Möller, P.; Dulski, P.; Pekdeger, A. Positive gadolinium anomaly in surface water and ground water of the urban area Berlin, Germany. *Chemie der erde-Geochemistry* **2005**, *65*, 167–189. [CrossRef]
154. Kümmerer, K.; Helmers, E. Hospital effluents as a source of gadolinium in the aquatic environment. *Environ. Sci. Technol.* **2000**, *34*, 573–577. [CrossRef]
155. Brünjes, R.; Bichler, A.; Hoehn, P.; Lange, F.T.; Brauch, H.-J.; Hofmann, T. Anthropogenic gadolinium as a transient tracer for investigating river bank filtration. *Sci. Total Environ.* **2016**, *571*, 1432–1440. [CrossRef]
156. Verplanck, P.L.; Taylor, H.E.; Nordstrom, D.K.; Barber, L.B. Aqueous stability of gadolinium in surface waters receiving sewage treatment plant effluent, Boulder Creek, Colorado. *Environ. Sci. Technol.* **2005**, *39*, 6923–6929. [CrossRef]

157. Merschel, G.; Bau, M.; Baldewein, L.; Dantas, E.L.; Walde, D.; Bühn, B. Tracing and tracking wastewater-derived substances in freshwater lakes and reservoirs: Anthropogenic gadolinium and geogenic REEs in Lake Paranoá, Brasilia. *Comptes Rendus Geosci.* **2015**, *347*, 284–293. [CrossRef]
158. Kümmerer, K. *Pharmaceuticals in the Environment: Sources, Fate, Effects and Risks*; Springer Science & Business Media: Berlin, Germany, 2008.
159. Verplanck, P.L.; Furlong, E.T.; Gray, J.L.; Phillips, P.J.; Wolf, R.E.; Esposito, K. Evaluating the behavior of gadolinium and other rare earth elements through large metropolitan sewage treatment plants. *Environ. Sci. Technol.* **2010**, *44*, 3876–3882. [CrossRef]
160. OECD. Magnetic Resonance Imaging (MRI) Exams. Available online: https://doi.org/10.1787/1d89353f-en (accessed on 7 December 2018).
161. Censi, P.; Zuddas, P.; Randazzo, L.; Saiano, F.; Mazzola, S.; Arico, P.; Cuttitta, A.; Punturo, R. Influence of dissolved organic matter on rare earth elements and yttrium distributions in coastal waters. *Chem. Ecol.* **2010**, *26*, 123–135. [CrossRef]
162. Möller, P.; Dulski, P.; Bau, M.; Knappe, A.; Pekdeger, A.; Sommer-von Jarmersted, C. Anthropogenic gadolinium as a conservative tracer in hydrology. *J. Geochem. Explor.* **2000**, *69*, 409–414. [CrossRef]
163. Rabiet, M.; Brissaud, F.; Seidel, J.; Pistre, S.; Elbaz-Poulichet, F. Positive gadolinium anomalies in wastewater treatment plant effluents and aquatic environment in the Hérault watershed (South France). *Chemosphere* **2009**, *75*, 1057–1064. [CrossRef] [PubMed]
164. Bichler, A.; Muellegger, C.; Brünjes, R.; Hofmann, T. Quantification of river water infiltration in shallow aquifers using acesulfame and anthropogenic gadolinium. *Hydrol. Process.* **2016**, *30*, 1742–1756. [CrossRef]
165. Sherry, A.D.; Caravan, P.; Lenkinski, R.E. Primer on gadolinium chemistry. *J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Reson. Med.* **2009**, *30*, 1240–1248. [CrossRef] [PubMed]
166. Shellock, F.G.; Spinazzi, A. MRI safety update 2008: Part 1, MRI contrast agents and nephrogenic systemic fibrosis. *Am. J. Roentgenol.* **2008**, *191*, 1129–1139. [CrossRef] [PubMed]
167. Mundim, J.S.; Lorena, S.d.C.; Abensur, H.; Elias, R.M.; Moysés, R.M.A.; MARTINS, M.C.; CASTRO, J.E.R. Fibrose sistêmica nefrogênica: Uma complicação grave do uso do gadolínio em pacientes com insuficiência renal. *Rev. Assoc. Med. Bras.* **2009**, *55*, 220–225. [CrossRef]
168. Ergün, I.; Keven, K.; Uruc, I.; Ekmekci, Y.; Canbakan, B.; Erden, I.; Karatan, O. The safety of gadolinium in patients with stage 3 and 4 renal failure. *Nephrol. Dial. Transplant.* **2005**, *21*, 697–700. [CrossRef]
169. Marckmann, P.; Skov, L.; Rossen, K.; Heaf, J.G.; Thomsen, H.S. Case-control study of gadodiamide-related nephrogenic systemic fibrosis. *Nephrol. Dial. Transplant.* **2007**, *22*, 3174–3178. [CrossRef]
170. Thakral, C.; Alhariri, J.; Abraham, J.L. Long-term retention of gadolinium in tissues from nephrogenic systemic fibrosis patient after multiple gadolinium-enhanced MRI scans: Case report and implications. *Contrast Media Mol. Imaging* **2007**, *2*, 199–205. [CrossRef]
171. Hasdenteufel, F.; Luyasu, S.; Renaudin, J.-M.; Paquay, J.-L.; Carbutti, G.; Beaudouin, E.; Moneret-Vautrin, D.A.; Kanny, G. Anaphylactic shock after first exposure to gadoterate meglumine: Two case reports documented by positive allergy assessment. *J. Allergy Clin. Immunol.* **2008**, *121*, 527–528. [CrossRef]
172. Kay, J. Nephrogenic systemic fibrosis: A gadolinium-associated fibrosing disorder in patients with renal dysfunction. *Ann. Rheum. Dis.* **2008**, *67*, iii66–iii69. [CrossRef] [PubMed]
173. White, G.W.; Gibby, W.A.; Tweedle, M.F. Comparison of Gd (DTPA-BMA)(Omniscan) versus Gd (HP-DO3A)(ProHance) relative to gadolinium retention in human bone tissue by inductively coupled plasma mass spectroscopy. *Investig. Radiol.* **2006**, *41*, 272–278. [CrossRef] [PubMed]
174. Darrah, T.H.; Prutsman-Pfeiffer, J.J.; Poreda, R.J.; Campbell, M.E.; Hauschka, P.V.; Hannigan, R.E. Incorporation of excess gadolinium into human bone from medical contrast agents. *Metallomics* **2009**, *1*, 479–488. [CrossRef] [PubMed]
175. Thakral, C.; Abraham, J.L. Gadolinium-induced nephrogenic systemic fibrosis is associated with insoluble Gd deposits in tissues: In vivo transmetallation confirmed by microanalysis. *J. Cutan. Pathol.* **2009**, *36*, 1244–1254. [CrossRef] [PubMed]
176. US Food and Drug Administration. *FDA Drug Safety Communication: FDA Evaluating the Risk of Brain Deposits with Repeated Use of Gadolinium-Based Contrast Agents for Magnetic Resonance Imaging (MRI)*; US Food and Drug Administration: Silver Spring, MD, USA, 2015.
177. Rogosnitzky, M.; Branch, S. Gadolinium-based contrast agent toxicity: A review of known and proposed mechanisms. *Biometals* **2016**, *29*, 365–376. [CrossRef] [PubMed]

178. Cowper, S.E.; Robin, H.S.; Steinberg, S.M.; Su, L.D.; Gupta, S.; LeBoit, P.E. Scleromyxoedema-like cutaneous diseases in renal-dialysis patients. *Lancet* **2000**, *356*, 1000–1001. [CrossRef]
179. Swaminathan, S.; High, W.; Ranville, J.; Horn, T.; Hiatt, K.; Thomas, M.; Brown, H.; Shah, S. Cardiac and vascular metal deposition with high mortality in nephrogenic systemic fibrosis. *Kidney Int.* **2008**, *73*, 1413–1418. [CrossRef]
180. Idee, J.-M.; Port, M.; Medina, C.; Lancelot, E.; Fayoux, E.; Ballet, S.; Corot, C. Possible involvement of gadolinium chelates in the pathophysiology of nephrogenic systemic fibrosis: A critical review. *Toxicology* **2008**, *248*, 77–88. [CrossRef]
181. Larson, K.N.; Gagnon, A.L.; Darling, M.D.; Patterson, J.W.; Cropley, T.G. Nephrogenic systemic fibrosis manifesting a decade after exposure to gadolinium. *JAMA Dermatol.* **2015**, *151*, 1117–1120. [CrossRef]
182. Sieber, M.A.; Lengsfeld, P.; Frenzel, T.; Golfier, S.; Schmitt-Willich, H.; Siegmund, F.; Walter, J.; Weinmann, H.-J.; Pietsch, H. Preclinical investigation to compare different gadolinium-based contrast agents regarding their propensity to release gadolinium in vivo and to trigger nephrogenic systemic fibrosis-like lesions. *Eur. Radiol.* **2008**, *18*, 2164–2173. [CrossRef] [PubMed]
183. Gathings, R.M.; Reddy, R.; Santa Cruz, D.; Brodell, R.T. Gadolinium-associated plaques: A new, distinctive clinical entity. *JAMA Dermatol.* **2015**, *151*, 316–319. [CrossRef] [PubMed]
184. Semelka, R.C.; Commander, C.W.; Jay, M.; Burke, L.M.; Ramalho, M. Presumed gadolinium toxicity in subjects with normal renal function: A report of 4 cases. *Investig. Radiol.* **2016**, *51*, 661–665. [CrossRef] [PubMed]
185. Semelka, R.C.; Ramalho, J.; Vakharia, A.; AlObaidy, M.; Burke, L.M.; Jay, M.; Ramalho, M. Gadolinium deposition disease: Initial description of a disease that has been around for a while. *Magn. Reson. Imaging* **2016**, *34*, 1383–1390. [CrossRef]
186. Burke, L.M.; Ramalho, M.; AlObaidy, M.; Chang, E.; Jay, M.; Semelka, R.C. Self-reported gadolinium toxicity: A survey of patients with chronic symptoms. *Magn. Reson. Imaging* **2016**, *34*, 1078–1080. [CrossRef] [PubMed]
187. Roberts, D.R.; Lindhorst, S.M.; Welsh, C.T.; Maravilla, K.R.; Herring, M.N.; Braun, K.A.; Thiers, B.H.; Davis, W.C. High levels of gadolinium deposition in the skin of a patient with normal renal function. *Investig. Radiol.* **2016**, *51*, 280–289. [CrossRef] [PubMed]
188. Semelka, R.C.; Ramalho, M.; AlObaidy, M.; Ramalho, J. Gadolinium in humans: A family of disorders. *Am. J. Roentgenol.* **2016**, *207*, 229–233. [CrossRef]
189. Abraham, J.L.; Chandra, S.; Thakral, C.; Abraham, J.M. SIMS imaging of gadolinium isotopes in tissue from Nephrogenic Systemic Fibrosis patients: Release of free Gd from magnetic resonance imaging (MRI) contrast agents. *Appl. Surf. Sci.* **2008**, *255*, 1181–1184. [CrossRef]
190. George, S.J.; Webb, S.M.; Abraham, J.L.; Cramer, S.P. Synchrotron X-ray analyses demonstrate phosphate-bound gadolinium in skin in nephrogenic systemic fibrosis. *Br. J. Dermatol.* **2010**, *163*, 1077–1081. [CrossRef]
191. Qin, Z.; Caruso, J.A.; Lai, B.; Matusch, A.; Becker, J.S. Trace metal imaging with high spatial resolution: Applications in biomedicine. *Metallomics* **2011**, *3*, 28–37. [CrossRef]
192. Jurowski, K.; Buszewski, B.; Piekoszewski, W. Bioanalytics in quantitive (bio) imaging/mapping of metallic elements in biological samples. *Crit. Rev. Anal. Chem.* **2015**, *45*, 334–347. [CrossRef] [PubMed]
193. Fingerhut, S.; Niehoff, A.-C.; Sperling, M.; Jeibmann, A.; Paulus, W.; Niederstadt, T.; Allkemper, T.; Heindel, W.; Holling, M.; Karst, U. Spatially resolved quantification of gadolinium deposited in the brain of a patient treated with gadolinium-based contrast agents. *J. Trace Elem. Med. Biol.* **2018**, *45*, 125–130. [CrossRef] [PubMed]
194. McDonald, R.J.; McDonald, J.S.; Kallmes, D.F.; Jentoft, M.E.; Murray, D.L.; Thielen, K.R.; Williamson, E.E.; Eckel, L.J. Intracranial gadolinium deposition after contrast-enhanced MR imaging. *Radiology* **2015**, *275*, 772–782. [CrossRef] [PubMed]
195. Kasahara, S.; Miki, Y.; Kanagaki, M.; Yamamoto, A.; Mori, N.; Sawada, T.; Taoka, T.; Okada, T.; Togashi, K. Hyperintense dentate nucleus on unenhanced T1-weighted MR images is associated with a history of brain irradiation. *Radiology* **2011**, *258*, 222–228. [CrossRef] [PubMed]
196. Xia, D.; Davis, R.L.; Crawford, J.A.; Abraham, J.L. Gadolinium released from MR contrast agents is deposited in brain tumors: In situ demonstration using scanning electron microscopy with energy dispersive X-ray spectroscopy. *Acta Radiol.* **2010**, *51*, 1126–1136. [CrossRef] [PubMed]

197. Kartamihardja, A.A.P.; Nakajima, T.; Kameo, S.; Koyama, H.; Tsushima, Y. Impact of impaired renal function on gadolinium retention after administration of gadolinium-based contrast agents in a mouse model. *Investig. Radiol.* **2016**, *51*, 655–660. [CrossRef]
198. Kanda, T.; Ishii, K.; Kawaguchi, H.; Kitajima, K.; Takenaka, D. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: Relationship with increasing cumulative dose of a gadolinium-based contrast material. *Radiology* **2013**, *270*, 834–841. [CrossRef]
199. Kanda, T.; Fukusato, T.; Matsuda, M.; Toyoda, K.; Oba, H.; Kotoku, J.i.; Haruyama, T.; Kitajima, K.; Furui, S. Gadolinium-based contrast agent accumulates in the brain even in subjects without severe renal dysfunction: Evaluation of autopsy brain specimens with inductively coupled plasma mass spectroscopy. *Radiology* **2015**, *276*, 228–232. [CrossRef]
200. Lingott, J.; Lindner, U.; Telgmann, L.; Esteban-Fernández, D.; Jakubowski, N.; Panne, U. Gadolinium-uptake by aquatic and terrestrial organisms-distribution determined by laser ablation inductively coupled plasma mass spectrometry. *Environ. Sci. Process. Impacts* **2016**, *18*, 200–207. [CrossRef]
201. Martino, C.; Bonaventura, R.; Byrne, M.; Roccheri, M.; Matranga, V. Effects of exposure to gadolinium on the development of geographically and phylogenetically distant sea urchins species. *Mar. Environ. Res.* **2017**, *128*, 98–106. [CrossRef]
202. Wáng, Y.-X.J.; Schroeder, J.; Siegmund, H.; Idée, J.-M.; Fretellier, N.; Jestin-Mayer, G.; Factor, C.; Deng, M.; Kang, W.; Morcos, S.K. Total gadolinium tissue deposition and skin structural findings following the administration of structurally different gadolinium chelates in healthy and ovariectomized female rats. *Quant. Imaging in Med. Surg.* **2015**, *5*, 534.
203. Lohrke, J.; Frisk, A.-L.; Frenzel, T.; Schöckel, L.; Rosenbruch, M.; Jost, G.; Lenhard, D.C.; Sieber, M.A.; Nischwitz, V.; Küppers, A. Histology and gadolinium distribution in the rodent brain after the administration of cumulative high doses of linear and macrocyclic gadolinium-based contrast agents. *Investig. Radiol.* **2017**, *52*, 324. [CrossRef] [PubMed]
204. McDonald, R.J.; McDonald, J.S.; Dai, D.; Schroeder, D.; Jentoft, M.E.; Murray, D.L.; Kadirvel, R.; Eckel, L.J.; Kallmes, D.F. Comparison of gadolinium concentrations within multiple rat organs after intravenous administration of linear versus macrocyclic gadolinium chelates. *Radiology* **2017**, *285*, 536–545. [CrossRef] [PubMed]
205. Bussi, S.; Coppo, A.; Botteron, C.; Fraimbault, V.; Fanizzi, A.; De Laurentiis, E.; Colombo Serra, S.; Kirchin, M.A.; Tedoldi, F.; Maisano, F. Differences in gadolinium retention after repeated injections of macrocyclic MR contrast agents to rats. *J. Magn. Reson. Imaging* **2018**, *47*, 746–752. [CrossRef] [PubMed]
206. Kanda, T.; Nakai, Y.; Oba, H.; Toyoda, K.; Kitajima, K.; Furui, S. Gadolinium deposition in the brain. *Magn. Reson. Imaging* **2016**, *34*, 1346–1350. [CrossRef] [PubMed]
207. Murata, N.; Gonzalez-Cuyar, L.F.; Murata, K.; Fligner, C.; Dills, R.; Hippe, D.; Maravilla, K.R. Macrocyclic and other non–group 1 gadolinium contrast agents deposit low levels of gadolinium in brain and bone tissue: Preliminary results from 9 patients with normal renal function. *Investig. Radiol.* **2016**, *51*, 447–453. [CrossRef] [PubMed]
208. Robert, P.; Violas, X.; Grand, S.; Lehericy, S.; Idée, J.-M.; Ballet, S.; Corot, C. Linear gadolinium-based contrast agents are associated with brain gadolinium retention in healthy rats. *Investig. Radiol.* **2016**, *51*, 73. [CrossRef]
209. Oh, K.Y.; Roberts, V.H.; Schabel, M.C.; Grove, K.L.; Woods, M.; Frias, A.E. Gadolinium chelate contrast material in pregnancy: Fetal biodistribution in the nonhuman primate. *Radiology* **2015**, *276*, 110–118. [CrossRef]
210. Zhimang, G.; Xiaorong, W.; Xueyuan, G.; Jing, C.; Liansheng, W.; Lemei, D.; Yijun, C. Effects of fulvic acid on the bioavailability of rare earth elements and GOT enzyme activity in wheat (Triticum aestivum). *Chemosphere* **2001**, *44*, 545–551. [CrossRef]
211. Braun, M.; Zavanyi, G.; Laczovics, A.; Berényi, E.; Szabó, S. Can aquatic macrophytes be biofilters for gadolinium based contrasting agents? *Water Res.* **2018**, *135*, 104–111. [CrossRef]
212. Yeo, A.R.; Yeo, M.E.; Flowers, T.J. The contribution of an apoplastic pathway to sodium uptake by rice roots in saline conditions. *J. Exp. Bot.* **1987**, *38*, 1141–1153. [CrossRef]
213. Nowack, B.; Schulin, R.; Robinson, B.H. Critical assessment of chelant-enhanced metal phytoextraction. *Environ. Sci. Technol.* **2006**, *40*, 5225–5232. [CrossRef] [PubMed]

214. Ternes, T.A. Occurrence of drugs in German sewage treatment plants and rivers1. *Water Res.* **1998**, *32*, 3245–3260. [CrossRef]
215. Roth, F.; Lessa, G.; Wild, C.; Kikuchi, R.; Naumann, M. Impacts of a high-discharge submarine sewage outfall on water quality in the coastal zone of Salvador (Bahia, Brazil). *Mar. Pollut. Bull.* **2016**, *106*, 43–48. [CrossRef] [PubMed]