

Geochemical baseline values determination and spatial distribution of trace elements in topsoils: An application in Sicily region (Italy)

Daniela Varrica^a, Federica Lo Medico^{a,*}, Marino Vetuschi Zuccolini^{b,c}, Marianna Miola^{b,c}, Maria Grazia Alaimo^a

^a Università degli studi di Palermo, DiSTeM, Palermo, Italy

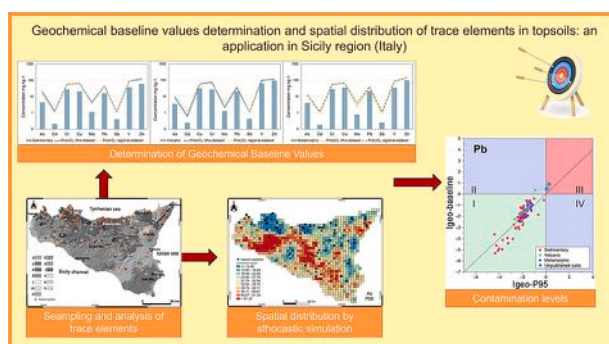
^b Università degli studi di Genova, DISTAV, Genova, Italy

^c CNR-IMATI, Istituto "E.Magenes", Genova, Italy

HIGHLIGHTS

- Statistical approach was used to define the geochemical baseline in the Sicily region.
- The geochemical baseline was determined using UTL95-95 values by ProUCL 5.1 software.
- Regional baseline value for each trace elements are suitable for the Sicily region.
- Geochemical maps using geostatistical method provided values in unsampled areas.
- Baseline values provide a statistical way to assess contamination in topsoils.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Bo Gao

Keywords:

Topsoil
Trace elements
Baseline value
Geochemical maps
Contamination indices

ABSTRACT

This study proposed an approach to determine the geochemical baseline values in topsoils. The chosen study area is Sicily (Italy), a region characterized by significant lithological heterogeneity. Eighty-three topsoil samples were collected at several sites away from potential anthropogenic pollution sources. The concentrations of potentially toxic elements (As, Cd, Cr, Cu, Mo, Pb, Sb, V, and Zn) were quantified via inductively coupled plasma (ICP-MS). The elements showed median concentrations in the range $68.8\text{--}0.23\ \mu\text{g g}^{-1}$ and the trend of abundance was: $\text{Zn} > \text{V} > \text{Cr} > \text{Cu} > \text{Pb} > \text{As} > \text{Mo} > \text{Sb} > \text{Cd}$. Regional geochemical baseline values for trace elements were determined using statistical methods ($\text{Me} \pm 2\text{MAD}$; P95 and UTL95–95). The use of UTL95–95 was found to be the most suitable, obtaining appropriate geochemical baseline values for the entire region, regardless of lithology. The spatial distribution of the elements was determined by stochastic simulations on a convex-concave boundary with a resolution of 5 km, obtaining detailed geochemical maps that predict the distribution of concentrations of each element even in unsampled areas. The results of this study provide a methodology for a more correct assessment of the environmental contamination status of soils.

* Corresponding author.

E-mail address: federica.lomedico@unipa.it (F. Lo Medico).

<https://doi.org/10.1016/j.scitotenv.2024.176951>

Received 6 August 2024; Received in revised form 9 October 2024; Accepted 13 October 2024

Available online 18 October 2024

0048-9697/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Trace elements in soils are found as derived from natural sources related to processes such as weathering of rocks and volcanic eruptions, as well as from human activities such as mining, industrial processes, and the use of fertilizers and pesticides. PTEs, or Potentially Toxic Elements (Pouret and Hursthouse, 2019), occur naturally in the soil at trace levels. When these elements accumulate in soils, they can be taken up by plants, entering the food chain and potentially impacting human health (He et al., 2005; Karn et al., 2021; Cicchella et al., 2022; Varrica et al., 2022).

The availability of PTEs in soils depends on many factors, such as the mineralogical composition of the bedrock (Cabral Pinto et al., 2017), environmental conditions (Bini et al., 2011), and pedogenic processes. Several studies have demonstrated that soil properties significantly influence the background values and behavior of potentially toxic elements in soils (Li et al., 2022; Xu et al., 2022). These properties include pH, clay content, iron and aluminum oxides, organic matter (Fernandes et al., 2018; Ramos-Miras et al., 2014; Souza et al., 2017; Zhang et al., 2019), mineralogy, and the elemental composition of the soil (de Lima et al., 2020). The determination of geochemical background values is necessary to quantify soil pollution status and assess environmental and human health risks. This is indeed a complex and critical task in environmental geochemistry, essential for attempt to distinguish between natural and anthropogenic concentrations of PTEs in environmental samples (Matschullat et al., 2000). Establishing the natural background levels of heavy metals in soils, which refers to the geochemical composition of virgin soils unaffected by human activities, is indeed challenging. The difficulty arises primarily because atmospheric deposition can introduce additional trace elements into soils, thereby masking their natural geochemical signatures (Cicchella et al., 2005; Albanese et al., 2007). In light of this, it is crucial to introduce the term of a geochemical baseline presented in 1993 at the International Geological Correlation Program as the Global Geochemical Baselines (Salminen and Tarvainen, 1997; Salminen and Gregorauskiene, 2000; Galan et al., 2008), refers to the natural variation in the concentration of an element in the surface environment, at a determined place and time (Santos-Francés et al., 2017). This concept includes the geogenic natural content (background level) and the anthropogenic contribution to soils (Salminen and Gregorauskiene, 2000; Cicchella et al., 2005; Frattini et al., 2006; Albanese et al., 2007; Galan et al., 2008). Calculating the geochemical baseline is indeed a more practical approach for assessing soil trace element concentrations, in the context of the current global environmental impact, and provide guidelines and quality standards in legislation and policy-making environment (Chen et al., 1999; Facchinelli et al., 2001; Wei and Wen, 2012; Sappa et al., 2020). Despite its importance, there is currently no accepted methodology that determines the background and baseline values, therefore various approaches are used depending on the specific context and objectives of the study (Matschullat et al., 2000; Reimann and Garrett, 2005; Gałuszka and Migaszewski, 2011). An overview of the methodologies can be summarized shortly in direct, indirect, and integrated approaches (Dung et al., 2013; Kicińska and Turek, 2017). In direct methods, also known as geochemical methods, the geochemical background value is determined by analyzing samples taken in pristine areas before the 19th century or from areas presumed to be minimally affected by human activities (Baize and Sterckeman, 2001; Gałuszka, 2007). The indirect methodologies use a statistical approach (the 95th percentile, Robust Statistics, Cumulative Frequency Distribution, and more). The integrated method combines direct and indirect methods, statistically analyzing samples from non-contaminated areas or characterized by low contamination. This produces better statistical results which, combined with geological knowledge of the area, make this approach more reliable (Gałuszka, 2007). Geostatistical methods are considered integrated methods that involve the use of techniques such as kriging and spatial interpolation to predict background values in regions based on sampled data.

Environmental geochemical mapping has indeed gained significant importance in recent years, in distinguishing anthropogenic pollution from natural background values (Plant et al., 2001).

The choice of the Sicily region, considered as a pilot site, was determined by the lack of analytical data on the current environmental conditions of superficial soils. Therefore, the methodological approach used represents the first attempt to define the regional geochemical baseline through (1) defining the concentrations of 9 trace elements (As, Cd, Cr, Cu, Mo, Pb, Sb, V, Zn), designing a comprehensive sampling plan covering the different lithological units across Sicily; (2) propose an approach for determining geochemical baseline values in topsoils in areas characterized by high lithological heterogeneity; (3) integrate the geochemical data from laboratory analyses using geostatistical methods to interpolate the data and generate individual maps for each trace element (As, Cd, Cr, Cu, Mo, Pb, Sb, V, Zn); (4) define contamination levels using the established baseline values and compare them with calculated concentrations.

2. Materials and methods

2.1. Study area

The study area is Sicily region (Italy), an island in the Mediterranean Sea (Fig.1). According to the Köppen and Geiger classification (Koppek et al., 2006), the study area is characterized by a Mediterranean climate. The average annual temperature ranges from 11 °C to 20 °C, and the total annual precipitation varies from an average annual value of 385 mm to 1192 mm (SIAS- Sicilian region: http://www.sias.regione.sicilia.it/pdf/Climatologia_sicilia.pdf). Sicily's complex geological framework is characterized by three main structural units due to the presence of varying lithospheric parts (Fig. 1). Facing the Tyrrhenian Basin, to the north, the emerging Siculo-Maghrebian chain is outlined, whose western edge, originating from the deformation of the Meso-Cenozoic Northern African continental margin, is made up of several palaeogeographic settings in which carbonate deposits prevail (Doglioni et al., 1999).

The eastern side of the region, specifically the Nebrodi and Peloritani mountains, is composed of clayey sandstone and metamorphic rocks. This area represents the segment of the geological chain that formed above the northward-dipping African plate. The South-Eastern part of Sicily is occupied by the Hyblean Plateau, which forms part of the undeformed foreland of the African Plate. This plateau is primarily composed of carbonates and clastic rocks, with some interbedded basic volcanites. The central area of the island consists of deformed syn- to post-orogenic marine deposits, including the Gessoso-Solfifera Formation, sandy calcarenite, and marly limestone. The Mount Etna volcanic region extends between the northern metamorphic complex and the south-eastern carbonate Hyblean Plateau. The volcanic rocks of Etna range from alkali basalt to trachyte, although most of the lavas are of hawaiitic composition (Liguori and Brucculeri, 2004).

2.2. Soil sample collection and analytical method

A total of 83 topsoil samples (depth of 0–20 cm) were collected far from anthropogenic sources (i.e. extra-urban roads, urban areas, industrial areas, cultivated land) in the Sicilian territory. The sampling plan design envisaged the total coverage of the different lithological units of Sicily by collecting 65 soils in the sedimentary substrate (SED), 8 in the volcanic substrate (VOLC), and 10 in the metamorphic substrate (MET). The distribution of the samples is shown in Fig. 1. The samples were collected using a plastic scoop, stored in PVC packages, and transported to the laboratory. They were dried at 105 °C in a stove (AG System, mod. G-therm) overnight; stones and plant materials were removed, and each sample was sieved using 500 µm nylon mesh, then milled in agate mortar. Soil analysis was carried out at an accredited laboratory "Activation Laboratories Ltd. (Ontario, Canada)". The topsoil samples were digested with a mixture of 3 ml hydrochloric acid and 1 ml

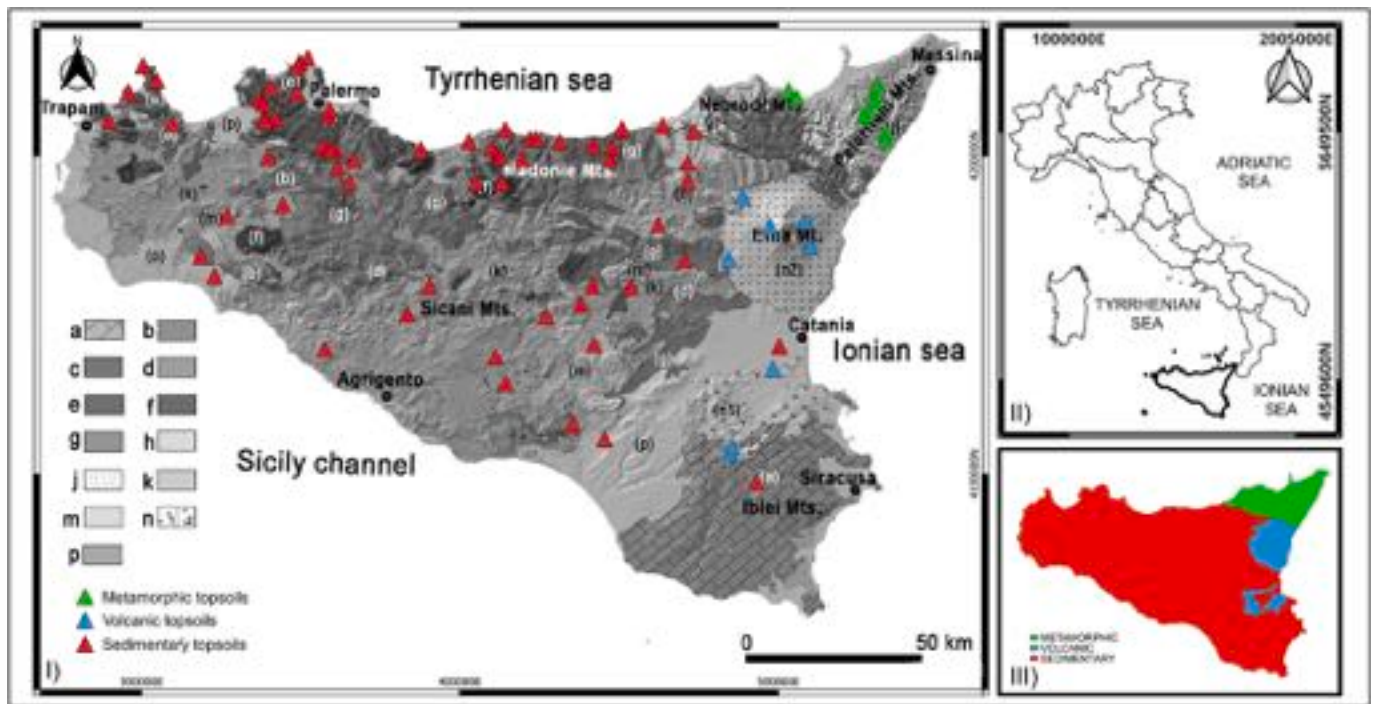


Fig. 1. (I) Topsoil samples (red: topsoil from sedimentary bedrock; blue: volcanic bedrock; green: metamorphic bedrock) and main geological domains of Sicily: (a) Hyblean units; (b) shelf to pelagic carbonate units; (c) shelf to deep-water carbonate units; (d) deep-water carbonate (Sicilian units); (e) shelf carbonate (Panormide) units; (f) slope to deep-water (Imerese-Panormide) units; (g) Miocene Flysch; (h) Sicilide Units; (j) Calabrian-Peloritan units; (k) Miocene–Pliocene syntectonic deposits; (m) Plio-Pleistocene syntectonic deposits; (n) Plio-Quaternary volcanic rocks; (p) Pleistocene deposits (modified from [Catalano and Sulli, 2011](#)); (II) geographical context of study area; (III) main lithological domains of the study area.

of nitric acid. The Activation Laboratories has automated the digestion method using a microprocessor designed hotbox to accurately reproduce digestion conditions every time. The samples were analyzed for Al, As, Cd, Cr, Cu, Mo, Pb, Sb, V, and Zn by ICP-MS. The OREAS-45d (anomalous ferruginous soil), OREAS-922 (low grade copper ore), OREAS-907 (copper-gold oxide ore), OREAS-263 (gold oxide), OREAS-130 (sedex type Zn-Pb-Ag ore), OREAS-521 (iron oxide copper-gold ore), OREAS-620 (volcanic hosted massive sulphide Zn-Pb-Cu-Ag-Au ore), OREAS-610 (High Sulphidation Epithermal Au-Cu-Ag Ore) standard reference materials were employed from Activation Laboratories Ltd. The metal recovery rates of certified elements in the reference standard materials were all >92%. Analytical precision, estimated from duplicate analyses every tenth sample was in the range of 15–22% for all analyzed elements. The mineralogical characterization of the soils was determined semi-quantitatively by X-ray powder diffraction (XRD) using a Philips PW14 1373 with Cu-K α radiation of the Department of Earth and Marine Sciences (DiSTeM), University of Palermo.

The measurement of pH values was performed potentiometrically on the soil suspensions obtained by adding 25 ml of potassium chloride 1 M to 10 g of soil (dry weight). The solution was kept agitated for at least two hours and then allowed to sit for 30 min.

2.3. Statistical analysis

The obtained results were analyzed using a multivariate statistical approach with the software XLSTAT (version 2023.3.0), R, and Rstudio (R Core Team, 2021). The data were explored using statistical parameters of central tendency (mean, median), dispersion (standard deviation), and shape (kurtosis coefficient, skewness index). The normality of data distributions was assessed using the Shapiro-Wilk test (W). The results indicate that the distribution of the analyzed elements is non-normal. All tests were considered significant at $p < 0.05$. Among the dispersion parameters, the robust coefficient of variation (%rCV) was used ([Ambrosino et al., 2023](#)). The robust coefficient of variation value

indicates the dispersion of the element that does not follow a normal Gaussian distribution. This was calculated as follows: $\%rCV = MAD / Median \times 100$, where MAD is the Median Absolute Deviation. To investigate the relationships between trace elements and identify possible distinct groups within the data, a principal component analysis (PCA) with varimax rotation was performed. To define relationships between a couple of variables Spearman's correlation coefficients (R-value) were used. The non-parametric Mann-Whitney test (U) was employed to compare differences between two or more independent groups.

2.4. Geochemical baseline, spatial distribution and contamination indices

The methodological approach proposed for the determination of the geochemical baseline values and an estimate of the regional elemental distribution through stochastic simulations is shown in [Fig. 2](#).

2.4.1. Geochemical baseline

Sicily region is characterized by significant differences in the soil chemical composition, resulting from the heterogeneity of the bedrock: sedimentary, volcanic, and metamorphic. Therefore, we also decided to segment the regional dataset into three sub-datasets: sedimentary (SED), volcanic (VOLC), and metamorphic (MET). Two geochemical baseline values are proposed for each trace element investigated: one regional value, determined from the regional dataset, and one lithological value defined for each sub-dataset ([Santos-Francés et al., 2017](#)).

Geochemical baseline values for trace elements were determined using statistical methods:

- $Me \pm 2MAD$ (Me: Median; MAD: Median Absolute Deviation), where the baseline value is determined as the upper limit of the interval ([Kicińska and Turek, 2017](#); [Reimann et al., 2005](#)).
- P95: Establishing baseline values at the 95th percentile involves identifying thresholds that separate most of the data from the upper tail of the distribution, representing low probable outcomes.

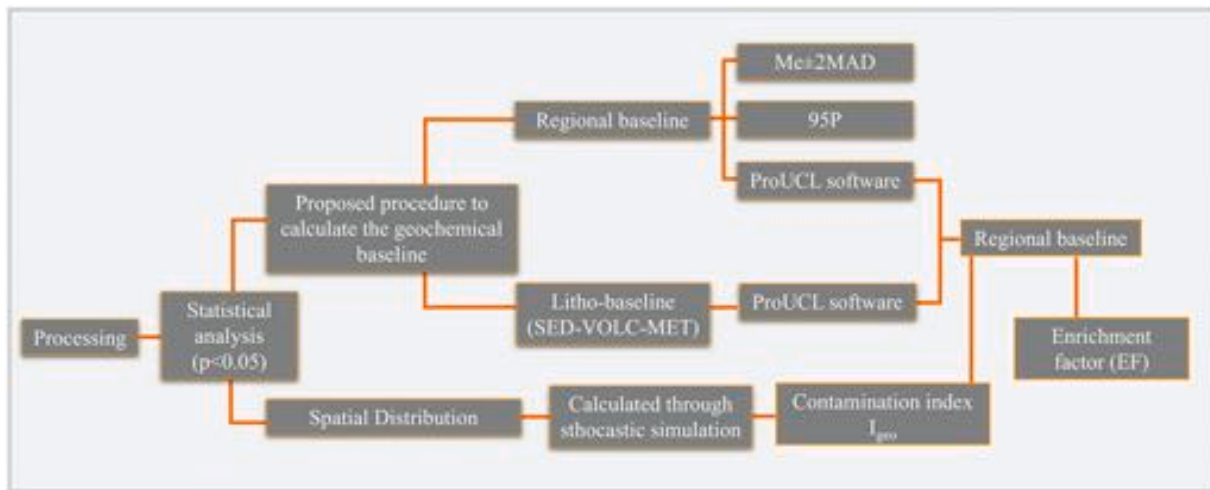


Fig. 2. Project work plan.

- UTL95–95 BCA Bootstrap (SNPA, 2018): The baseline value is calculated using the ProUCL 5.1 software (Singh and Maichle, 2015, <https://www.epa.gov/land-research/proucl-software>), which provides results based on the type of distribution. In the case of non-normal distributions, the non-parametric UTL95–95 value is used. UTL95–95 corresponds to the Upper one-sided Tolerance Limit at 95 % (UTL) with a 95 % coverage (UTL95–95). The BCA bootstrap method (Efron and Tibshirani, 1993; Manly, 1997) is used to estimate the upper tolerance limit (UTL), correct bias, and accelerate the confidence interval. This method creates multiple bootstrap samples from the original data and calculates the upper limit for each sample. Subsequently, a bias correction is applied to adjust the results to eliminate any systematic distortions in the estimates. This is done by comparing the values obtained from the bootstrap samples with those of the original sample. Finally, the acceleration of the data is an additional adjustment that considers the variability of the data and helps make the confidence intervals more accurate.

2.4.2. Spatial distribution

The regional dataset, composed of 83 samples collected on a heterogeneous geological area, shows a clustered pattern featured also by redundant sampling sites. To generate elemental distributions a deterministic algorithm based on hard data, considered error-free, would have as a result a distribution very continuous as with B-spline, low-pass filtered or kriging family algorithms (Cressie, 1990) and prone to artifacts as in inverse distance weighting. Isolated samples would receive an excessive weight in computation leading to widespread, high or low, spatial bulging anomalies. To take into account an expected high frequency and short-scale variability in element content in soils, the deterministic rules were replaced by a set of covariance spatial laws defined through omnidirectional variograms (Goovaerts, 1997). The spatial distributions have been calculated through stochastic simulations (Deutsch and Journel, 1998; Gómez-Hernández and Srivastava, 2021) on a convex-concave boundary at a resolution of 5 km. Due to the small number of samples used in variography, the stochasticity could not follow the same segmentation as for the baseline calculation, then maps were computed for the whole domain. Each of the 1067 cells used to discretise Sicily was calculated by dividing it by four sub-cells for 50 simulations. Biases due to heterogeneous sampling scheme, and asymmetric probability distribution function, were reduced through simulations performed on the normal space (Deutsch, 1989) by anamorphosis to normal standard distribution and then back-transformed to original compositional space. The use of this geostatistical technique allows the estimation of local uncertainty of the estimated content values in each cell of the computational grid.

2.4.3. Contamination indices

The Enrichment Factor (EF) is defined by Yongming et al. (2006) as Eq. (1):

$$EF = \frac{\left(\frac{X}{Ref}\right)_{sample}}{\left(\frac{X}{Ref}\right)_{baseline}} \quad (1)$$

where X represents the element of interest and Ref the reference element respectively in the sample (numerator) and in the calculated regional baseline (denominator). There are no universal fixed rules for the choice of reference element, except that it must be immobile and almost exclusively of crustal origin. In calculating the enrichment factor, Al, Fe, Sc, or Ti are the most used elements as reference. In this study, we used aluminum as a reference element based on these considerations: 1) it is high natural abundance, and 2) it is easily determined by conventional techniques.

The contamination categories based on Enrichment Factor (EF) were (Yongming et al., 2006): EF < 2 Deficiency to minimal enrichment; 2 < EF < 5 Moderate enrichment; 5 < EF < 20 Significant enrichment; 20 < EF < 40 Very high enrichment; EF > 40 Extremely high enrichment.

The Geo-accumulation Index (I_{geo}) is used to assess the net accumulation of heavy metals or other chemical elements in sediments or soils over time. The I_{geo} is calculated from Eq. (2) (Suresh et al., 2012):

$$I_{geo} = \text{Log}_2[C_i / (1.5B_i)] \quad (2)$$

where C_i is the measured concentration of the i -th metal examined in the topsoil sample, and B_i is the baseline level of the i -th metal. Factor 1.5 was used to correct possible variations in the background values of a particular metal in the environment (Santos-Francés et al., 2017). According to Müller (1969), the Geo-accumulation index (I_{geo}) consists of 7 classes: uncontaminated ($I_{geo} \leq 0$), uncontaminated to moderately contaminated ($0 < I_{geo} \leq 1$), moderately contaminated ($1 < I_{geo} \leq 2$), moderately to heavily contaminated ($2 < I_{geo} \leq 3$), heavily contaminated ($3 < I_{geo} \leq 4$), heavily to extremely contaminated ($4 < I_{geo} \leq 5$) and extremely contaminated ($I_{geo} > 5$).

3. Results and discussion

3.1. Characterization of topsoils

The relative proportion of minerals recognized in the soil samples reflects the heterogeneity of the types of soils and the different geological environments of formation. In the samples associated with a

sedimentary bedrock, calcite and clay minerals, such as kaolinite, illite, smectite, and chlorite, were the prevailing minerals, followed by plagioclase and quartz. These mineral associations characterize the sedimentary rocks of the study area, which are predominantly carbonate and clayey (Dongarrà and Ferla, 1982; Barbera et al., 2011; Saccà et al., 2011). In the soil samples collected in volcanic bedrock, the plagioclase is identified as the main component, followed by quartz and clinopyroxene, in agreement with the literature data, which identify the same mineral associations in the rocks of the Iblean Plateau and the Etnean area (Peccerillo, 2005; Mollo et al., 2011). In the topsoils associated with a metamorphic bedrock, illite and plagioclase are predominantly present, with calcite and quartz in subordinate amounts, and muscovite and chlorite as alteration products. These mineral associations are typical of the metamorphic rocks characterizing the study area, such as phyllites, gneisses, amphibolites, mica schists, quartzites, and metabasalts (Bonardi et al., 1982; De Vivo, 1982; Barbera et al., 2009; CARG, 2011).

The pH values determined in the soil samples showed values between 5.9 and 7.2. Based on the USDA (2017) classification, soil samples that fall in the range of 5.6–6.0 are defined as moderately acidic. This category includes 15.4 % of the samples that were collected in areas characterized by clayey bedrock/marly clay. In the pH range 6.1–6.5, they are defined as weakly acidic soils and correspond with samples taken in volcanic and metamorphic areas (15.4 % of total samples). Finally, samples included in the range 6.6–7.3 are classified as neutral and involve soils that insist on carbonate rocks corresponding to 69.2 % of the total samples. According to the literature, the variations in soil pH are influenced by the parental materials (Reuter et al., 2008; Gruba and Socha, 2016). The sedimentary rocks produce neutral-alkaline soils, while crystalline bedrocks tend to produce acidic soils (Fabian et al., 2014).

In Table 1 are reported the main statistical parameters of the trace elements measured in the topsoil samples grouped as a single dataset. The results of the Shapiro-Wilk normality test ($p < 0.05$) suggest an asymmetric distribution for all elements. The calculation of the skewness index shows a right-handed skewness for all elements. The values of the kurtosis coefficient indicate platykurtic-type distributions for Cr, Cu, Pb, and Zn. The other elements show a value indicative of leptokurtic-type distributions. The elements showed median concentrations in the range 68.8–0.23 $\mu\text{g g}^{-1}$ and had the following order of abundance: $\text{Zn} > \text{V} > \text{Cr} > \text{Cu} > \text{Pb} > \text{As} > \text{Mo} > \text{Sb} > \text{Cd}$. To evaluate the variability of the data, it calculated the robust coefficient of variation (%rCV). In general, a lower %rCV value corresponds to less relative variation in its concentrations. In contrast, higher %rCV values suggest that the median absolute deviation (MAD) is relatively large compared to median values, indicating a greater degree of variation in element concentrations. The robust coefficient of variation calculated for As, Cd, Cr, Cu, Mo, Pb, Sb, V, and Zn in topsoil samples showed wide variability for all elements (>30 %) as defined by Ambrosino et al. (2023), highlighting the lithological heterogeneity of the study area. A Principal Component Analysis (PCA) was performed to identify the different lithological domains in the total dataset (Table 2). The PCA indicates that the first three factors can

Table 1

Statistics of trace element content in topsoil samples. Concentration data expressed in $\mu\text{g g}^{-1}$.

	Mean	Median	Range	Kurtosis	Skewness	%rCV
Al	20,399	19,400	68.0–91,006	6.56	2.08	31
As	7.48	4.50	0.17–113	45.2	6.24	42
Cd	0.44	0.23	0.04–4.30	21.3	4.02	35
Cr	30.5	28.0	3.64–94.0	1.06	0.97	46
Cu	31.3	25.7	3.91–130	3.87	1.81	53
Mo	1.54	1.16	0.40–12.5	26.1	4.54	34
Pb	20.4	16.5	1.50–111	12.4	2.78	44
Sb	0.55	0.40	0.07–7.75	63.9	7.56	45
V	44.9	40.0	1.00–238	8.49	2.23	50
Zn	68.5	68.8	0.50–251	2.26	1.05	42

Table 2

Principal Component Analysis (PCA) for the metals and metalloids. Factor loadings and communalities for the first three factor, after Varimax rotation. Only factor loading values >0.7 are considered. Expl.Var: explained variance, % Var: % of variance.

	Factor 1	Factor 2	Factor 3
As	−0.057	−0.247	0.813
Cd	−0.042	0.797	−0.015
Cr	0.622	−0.393	0.013
Cu	0.762	−0.104	−0.108
Mo	−0.193	0.652	−0.017
Pb	0.707	0.176	0.199
Sb	0.131	0.195	0.819
V	0.761	−0.306	−0.134
Zn	0.810	−0.129	0.193
Expl.Var	2.761	1.466	1.441
% Var	30.7	16.3	16

explain approximately 63 % of the variance. Factor 1, which shows high positive loadings of elements V, Cu, Zn, and Pb, accounts for 30.7 % of the total variance in the dataset. Due to the dominance of these elements, we can identify this factor with a volcanic-metamorphic environment. Previous studies in the Etna area and the Peloritani mountains confirm the association of these trace elements (De Vivo et al., 1993; Aiuppa et al., 2000). The dominant elements in factor 2 are As and Sb (16.3 % of the total variance). This factor is mainly attributed to poly-metallic mineralizations with As and Sb sulfides present in metamorphic rocks of Peloritani Mountains (De Vivo et al., 1993; Ferla and Omenetto, 2000; Dongarrà et al., 2009). The profile of factor 3, which accounts for a further 16 % of the remaining variance, is essentially determined by Cd, which can be attributed to a sedimentary environment. The presence of Cd in sedimentary rocks has been found in several studies (Rambeau et al., 2010; Kubier et al., 2019). Giacalone et al. (2005), in the study area observed a greater presence of this element in brown soil on limestone and flyshoids sequence.

A relation among the trace elements was evaluated by a Spearman correlation matrix. The elements mostly show positive correlations. A highly significant statistical correlation with $R > 0.60$ ($p < 0.05$) was found between five elemental associations as As–Cr, Cr–V, Cu–V–Zn, Pb–Zn, and V–Zn. The correlations found as shown in Fig. 3 provide information on geochemical behavior and environmental distribution of trace elements. To evaluate whether there are compositional differences for the elements studied between the three lithological groups that characterize the Sicily region, the regional dataset was slitted into three sub-datasets: sedimentary (SED), volcanic (VOLC), and metamorphic (MET). The statistical distributions related to these sub-datasets are shown in Fig. 4. Statistical significant differences were determined by a non-parametric Mann-Whitney test (U) ($p < 0.05$). Different letters above each boxplot indicated significant differences (Fig. 4). Statistically significant differences are detected for Cu, V and Zn when comparing the SED-VOLC datasets and Cu–Zn between the SED-MET datasets. Besides the Cu, the U test does not report statistically significant differences between the VOLC-MET dataset. The higher median Cu and V concentrations in the volcanic soil samples highlight the direct influence of pedogenic processes in the basaltic rocks surrounding the Etna volcano (Aiuppa et al., 2000). Elements such as Cr and Zn are more abundant in samples belonging to the metamorphic domain. Their presence is also associated with the sulfide mineralizations in the Peloritani mountains (Ferla and Omenetto, 2000; Dongarrà et al., 2009). Therefore, the content of trace elements is characteristic of the substratum present.

3.2. Determination of geochemical baseline values and distribution maps

The calculations to evaluate the regional geochemical baseline were performed using three non-parametric statistical methods to account for the non-normal distributions of elements in all datasets. The methods

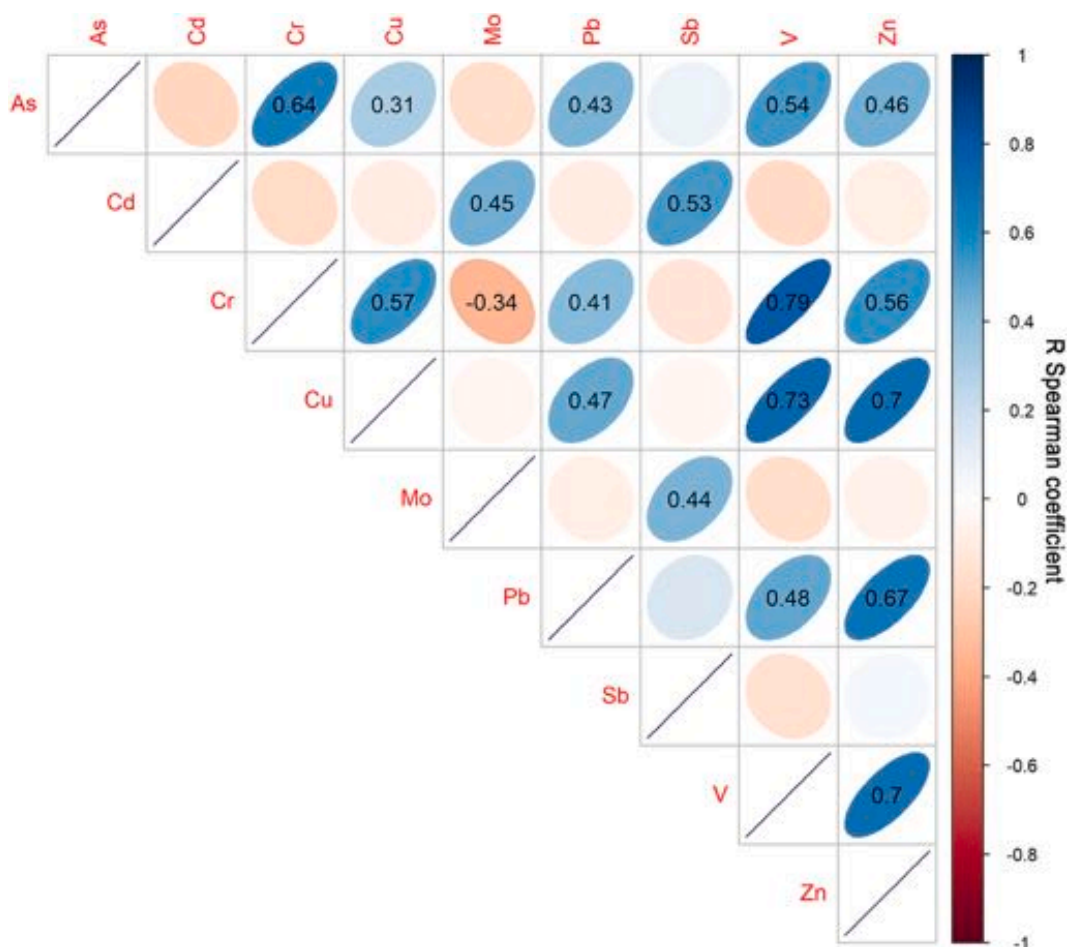


Fig. 3. Spearman correlation matrix of variables measured at the soil sample.

used were $Me \pm 2MAD$, the calculation of the 95th percentile, and the UTL95–95 value calculated using ProUCL 5.1 software. From the comparison of the methodologies to evaluate a regional geochemical baseline (Table 3), it is observed that the values determined with the MAD method generally appear more conservative than the other two methods. This observation is in full agreement with Reimann et al. (2005), who deduced that the $Me \pm 2MAD$ method tends to provide lower baseline values than other methods. The regional geochemical baseline calculated with the other two methods appears to be comparable for many elements. The 95th percentile is simpler but less robust to data distortions, while the UTL95–95 method with BCA bootstrap is more robust as it considers variability and bias in the data. Therefore, among the applied methods UTL95–95 is the most suitable to evaluate upper limits in quality control and environmental monitoring contexts. The concentration of elements exceeding the corresponding UTL95–95 values may indicate natural deposits of elevated metals and/or anthropogenic inputs (Crane, 2021). Additionally, UTL95–95 values can be used to monitor future trends in surface soils over time if sampling techniques, soil depth intervals, and methods for determining geochemical baseline values remain consistent (Crane, 2021). Considering this, the UTL 95–95 method was also used for determining the litho-baseline in the three sub-datasets (SED-VOLC-MET). Table 3 shows the regional baseline and litho-baseline geochemical values. The data comparison highlights agreement between the calculated values of the regional baseline and the litho-baseline in the sedimentary sub-dataset. At the same time, slight deviations are observed for some elements in the remaining sub-datasets. Especially, deviations for Cd and Mo in the VOLC dataset compared to the regional dataset of about -30% , for Mo in the MET dataset of about -22% , and for As in the SED dataset of

about $+25\%$ were calculated. It is therefore possible to consider the appropriate regional geochemical baseline for the entire region, regardless of lithology, while being aware that this may lead to underestimations or overestimations of these elements in the estimation of contamination indices.

In order to obtain the spatial distribution of PTE concentrations, and to suggest an alternative way of featuring the baseline with a local point of view, a set of geochemical maps was created. This has been done through geostatistical method that allowed us to predict the spatial distribution of the concentrations of each PTE in the unsampled areas in a discrete element, based on the values of the collected samples. The database used to calculate the geochemical numerical model consists of 83 samples, distributed within the regional territory according to a random scheme. The numerical distribution is based on a probabilistic approach based on stochastic simulations. The development of a regionally extended numerical tool is closely related to the quantification of baseline values. The applied numerical model reports the spatial statistical distributions of the analyzed elements (As, Cd, Cr, Cu, Mo, Pb, Sb, V, and Zn). As a result of the stochastic process over 50 simulations for each selected element, considering sub-cells subdivision, each cell of the spatial distribution model contains 200 values. Based on this dataset we are enabled to build a cumulative probability distribution function from where we extract the relevant statistical moments as Q1, mean, median, Q3 and 95th percentile (labeled as P95). P95, that is often chosen as a relevant threshold to identify extreme values as possible evidence of geochemical anomaly, is, appropriately plotted in Fig. 5. The representation of the elemental distribution shows a close relationship with the bedrock lithotypes.

The contamination status of the surface soil layer was assessed by

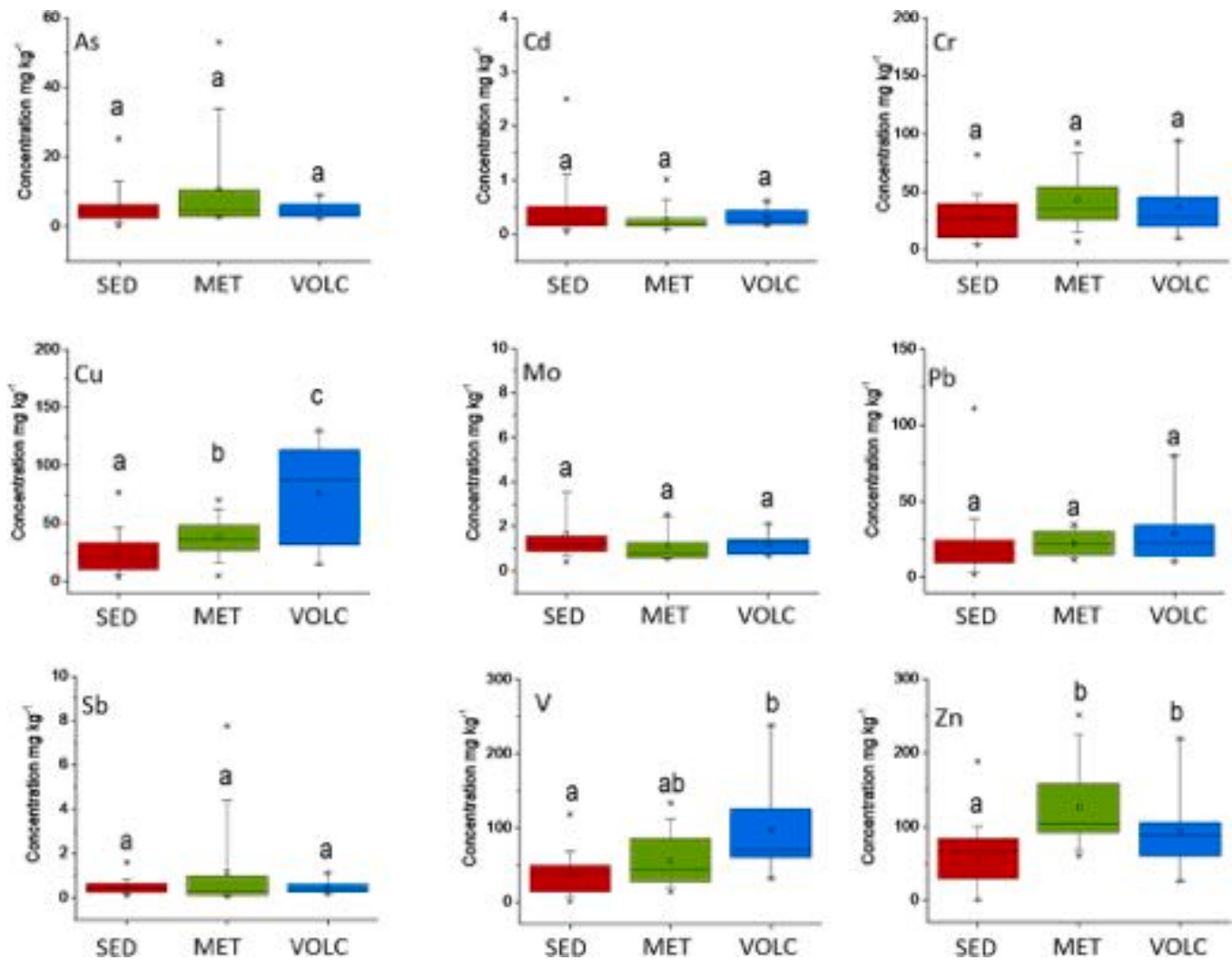


Fig. 4. Boxplots for each trace element distribution within the sub-dataset. SED: sedimentary soils; VOLC: volcanic soils and MET: metamorphic soils. Boxes delineate interquartile range (25–75 %) with an indication of the median (solid line) and mean value (small square inside the box). Whiskers indicate 10–90 % range. The asterisk symbols indicate the minimum and maximum values. The different letters upper each boxplot indicate the significant differences ($p < 0.05$). Data are expressed in $\mu\text{g g}^{-1}$.

Table 3

The MAD, P95, and UTL95–95 methods were employed to calculate the regional geochemical baseline values of trace elements. Additionally, the UTL95–95 method was used to determine the lithological baseline values of these trace elements.

	Regional geochemical baseline			Lithological baseline (UTL95–95)		
	MAD	P95	UTL95–95	Sedimentary	Metamorphic	Volcanic
Al	31,200	36,785	41,414	50,403	31,300	38,400
As	8.00	15.0	11.8	19.8	15.0	9.00
Cd	0.37	1.12	1.16	1.25	1.01	0.62
Cr	53.5	54.6	58.1	58.9	54.0	76.8
Cu	51.8	63.9	62.9	63.2	70.2	58
Mo	1.85	3.56	3.96	4.23	2.52	2.10
Pb	30.0	40.7	41.3	42.3	34.9	35.3
Sb	0.74	1.05	1.14	1.14	1.00	1.12
V	71.0	86.3	88.2	91.0	89.0	102
Zn	126	116	124	116.4	158	116

comparing the Enrichment Factor (EF) and the Geoaccumulation Index (I_{geo}). The classical formulation of the EF according to Zoller et al. (1974) envisaged the use of crustal concentrations as reference values. Among the critical issues of the use of the EF is the choice of the reference substrate (Reimann and de Caritat, 2000, 2005), since the composition of the earth's crust differs profoundly from the

representative composition of the rocks of the study area. Previous studies (Yongming et al., 2006; Sierra et al., 2015) have shown that using baseline values as a reference is a good alternative for the calculation and interpretation of the enrichment factor when the local rock/soil values for the study area are unknown. In this study, EF values were calculated for all samples using the regional geochemical baseline values

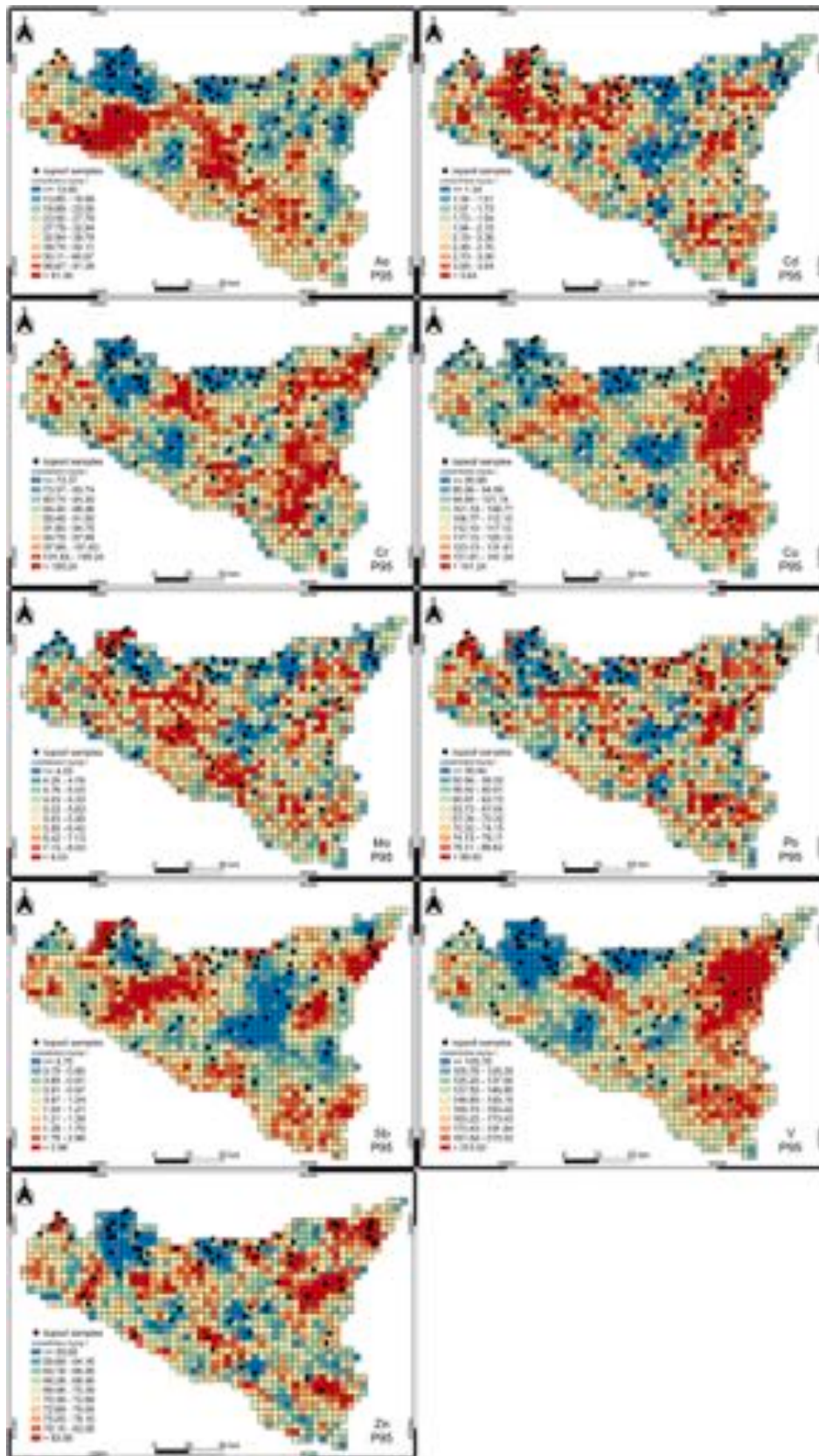


Fig. 5. Distribution map of PTEs. Data expressed in $\mu\text{g g}^{-1}$.

as a reference value.

In Fig. 6, box plots were used to synthetically represent the statistical distribution of the enrichment factor. A chromatic scale, from lightest to darkest, was used to highlight the increase in the enrichment factor. The enrichment factors distribution shows that median values calculated with the regional baseline are all < 2 and can be classified as “Minimally Enriched Deficiency” according to Yongming et al. (2006). Some elements (As, Cd, Mo, Pb, Sb, V, and Zn) show maximum values higher than 5. Based on the classification of Yongming et al. (2006) they should be considered “significant enrichment”. The anomaly found in these elements can be explained by the heterogeneity of the parent material that outcrops in the study area. In particular, Cd is influenced by flysch and carbonate rocks; as for the elements listed above, their enrichment is consequent to the presence of volcanic and metamorphic rocks.

The Geo-accumulation index (I_{geo}) is calculated for each element, using the determined regional baseline as B_i . An alternative evaluation of the I_{geo} , aimed at highlighting the geochemical imprint and compositional variability at the local scale of a single soil sample, can be performed considering a locally variable B_i , as suggested by the P95 distribution derived from probabilistic maps (Fig. 5). In Fig. 7 are compared the I_{geo} values of all soil samples labeled by their lithotype domain (SED, VOLC, MET), calculated with a locally variable B_i as P95 of the stochastic maps on the x-axis and with a constant regional baseline, B_i on the y-axis.

Each diagram is partitioned into four domains (I, II, III, IV) by the 0 value and bisected by the 1:1 identity line of $I_{geo-P95}$ versus $I_{geo-baseline}$ values which split each axis in the uncontaminated-contaminated domain. By comparing these two values, there is no misclassification of absolute values, as there is no “true” I_{geo} , but rather an assessment of how local effects on PTE content can affect the index relative to a regional baseline. The plots in Fig. 7 provide twofold information, one on the consistency between the results obtained by the two methods, and a second about the over- or under-estimation of one method by the other. In the green area marked I, the two I_{geo} 's predict the same condition of non-contaminated soils, while the red area marked III highlights the contamination, with all points below or above zero I_{geo} . In the blue areas, marked as II and IV, the comparison of the two values suggests if the local geological conditions can affect the metal content, particularly evaluating the distance of a point from the identity line and from the uncontaminated-contaminated limits. The behavior of the points cloud in each plot is mostly monotonic showing a trend close to

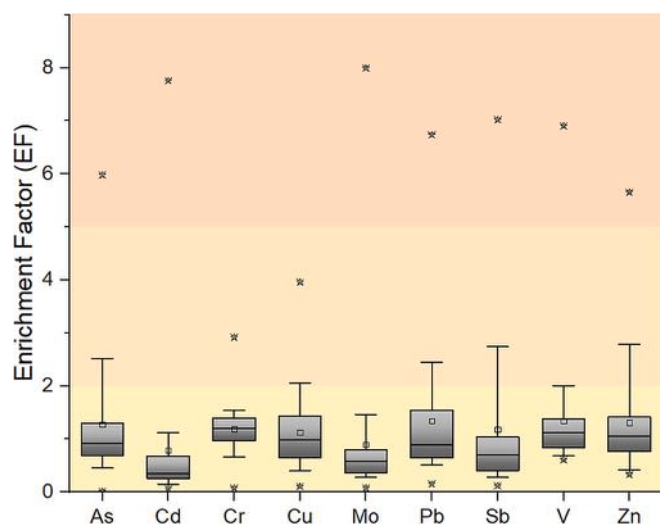


Fig. 6. Boxplots for Enrichment Factor (EF) value for each trace element. Boxes delineate interquartile range (25–75 %) with an indication of the median (solid line) and mean value (small square inside the box). Whiskers indicate 10–90 % range. The asterisk symbols indicate the minimum and maximum values.

the identity line apart from the arsenic, and with soils in sedimentary lithotypes that demonstrate generally lower I_{geo} magnitudes for both methods if compared with other lithotypes. The pattern of Sicily soil samples used in the paper, as shown in Fig. 7, proves quite robustly that media does not suffer from an anthropic bias. In the plots of As, Cd, and Sb very few samples fall in the III quadrant (contamination) highlighting the strong dependence on the parent rock. A cluster of volcanic samples for Cu classified as contaminated using a regional baseline, is considered differently acceptable on a probabilistic framework with $I_{geo} < 0$. The 83 samples used in the mapping and plotted in Fig. 7, stratified by lithotype, are used as a pattern background for comparison with 5 samples recently collected on metamorphic rocks in the Peloritani chain (unpublished, personal communication Federica Lo Medico). The point locations of these new samples in Fig. 7 are consistent with the previous dataset, with an overlap that, for Pb, indicates a good agreement with a regional trend. In fact, an underestimation of the $I_{geo-baseline}$ is contrasted with an overestimation with respect to $I_{geo-P95}$ for soil samples collected in sedimentary, volcanic, and metamorphic outcrops. Two samples are classified as slightly contaminated by both methods to the far right of the trend.

A comparative analysis of regional baseline values with respect to European threshold levels, such as those for agricultural soils (Reimann et al., 2014), the European ESDAC project (Panagos et al., 2022) and the Italian legislation (D. Lgs 152/ 2006), as illustrated in Table 4, clearly highlights that the regional approach offers more specific and meaningful indications than the general European conditions. The regional baseline values with respect to the Italian legislation data are all lower, while with respect to the threshold value ranges provided by the European Community and ESDAC, our data are close to the minimum values. Considering this, it can be concluded that regional values provide a more accurate indication of soil quality than the general European values, which may not adequately reflect the specific conditions of an unpolluted area and local geological characteristics. This approach allows a more targeted and reliable assessment of environmental contamination and supports a more effective land management.

4. Conclusion

Trace elements in soil can originate either from the natural weathering of rocks or from pollution resulting from human activities. Distinguishing between natural levels and anthropogenic inputs is essential, as these values vary depending on the region and the scale of the area being studied. Therefore, geochemical soil monitoring plays a key role in assessing the natural concentrations of heavy metals in soil, which are linked to parent materials and potential enrichment from human activities.

In this study the methodological approach used represents the first attempt to define the geochemical baseline in the Sicily region. The study area is characterized by a large geological heterogeneity and by the lack of data on the concentration of trace elements in surface soils. The values of the regional geochemical baseline were determined using a statistical approach by selecting the UTL95–95 value calculated using the ProUCL 5.1 software. The determination of these values can provide a more accurate indication in the determination of some environmental contamination indicators that allow to verify the possible origin of trace elements. The use of geochemical maps calculated with the geostatistical method provided a spatial distribution of PTE concentrations, highlighting how the lithology of the rock substrate governs their distribution and providing values also in areas not covered by sampling. A comparison between I_{geo} values based on stratified regional (statistical) and local (probabilistic) baselines can help to understand the need for deconvolution between natural and anthropogenic signals. This study effectively defined metal concentrations, established geochemical baseline values, created detailed geochemical maps and can provide valuable information to environmental databases to help in the management of soil contamination in Sicily. The inappropriate use of

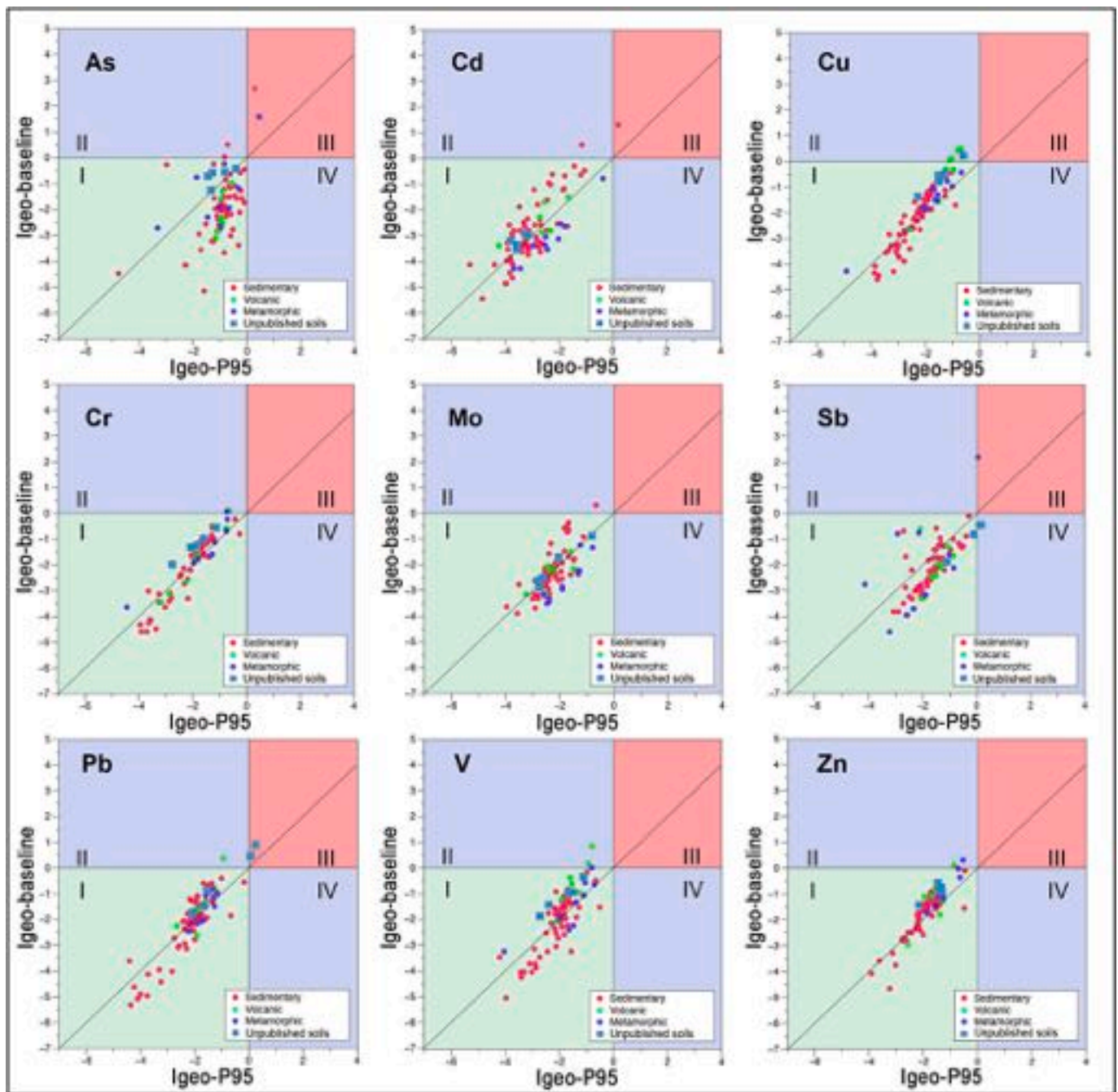


Fig. 7. The plots of Geo-accumulation indices ($I_{geo-P95}$ Vs $I_{geo-baseline}$). Quadrant I: uncontaminated ($I_{geo} \leq 0$); quadrant II: contaminated for $I_{geo-baseline}$ ($I_{geo} > 0$) and uncontaminated for $I_{geo-P95}$; quadrant III: contaminated ($I_{geo} > 0$) and quadrant IV: contaminated for $I_{geo-P95}$ ($I_{geo} > 0$) and uncontaminated for $I_{geo-baseline}$.

Table 4
Comparison of regional baseline values with to European threshold levels (Reimann et al., 2014), the European ESDAC project (Panagos et al., 2022) and the Italian legislation (Legislative Decree 152/2006); n.d.: not detected.

	Regional baseline	D.Lgs. 152/2006	ESDAC	Range EU
As	12	20	2.66–19.1	10.0–200
Cd	1	2	0.06–0.50	0.50–20.0
Cr	58	150	9.70–47.0	30.0–1000
Cu	63	120	7.80–35.9	40.0–1000
Mo	4	n.d.	n.d.	2.50–60.0
Pb	41	100	6.80–47.0	40.0–750
Sb	1	10	n.d.	10.0–50.0
V	88	90	n.d.	100–220
Zn	124	150	22.5–127	60.0–2500

normative values from environmental legislation of other countries or regions leads to a biased assessment of environmental impact, considering that the trace element content in soil is highly variable and strongly dependent on the parent rock.

Funding

This work was supported by MIUR FFR 2023 (M.G. Alaimo and D. Varrica).

CRedit authorship contribution statement

Daniela Varrica: Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis,

Data curation, Conceptualization. **Federica Lo Medico**: Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Data curation, Conceptualization. **Marino Vetuschi Zuccolini**: Writing – review & editing, Methodology, Formal analysis, Data curation. **Marianna Miola**: Formal analysis, Writing - review & editing. **Maria Grazia Alaimo**: Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

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.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- Aiuppa, A., Allard, P., D'Alessandro, W., Michel, A., Parello, F., Treuil, M., Valenza, M., 2000. Mobility and fluxes of major, minor and trace metals during basalt weathering at Mt. Etna volcano (Sicily). *Geochim. Cosmochim. Acta* 64, 1827–1841.
- 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, 21–34.
- Ambrosino, M., El-Saadani, Z., Abu Khatita, A., Mingqi, W., Palarea-Albaladejo, J., Cicchella, D., 2023. Geochemical speciation, ecological risk and assessment of main sources of potentially toxic elements (PTEs) in stream sediments from Nile River in Egypt. *Water* 15 (13), 2308.
- Baize, D., Sterckeman, T., 2001. Of the necessity of knowledge of the natural pedochemical background content in the evaluation of the contamination of soils by trace elements. *Sci. Total Environ.* 264, 127–139.
- Barbera, G., Lo Giudice, A., Mazzoleni, P., Pappalardo, A., 2009. Combined statistical and petrological analysis of provenance and diagenetic history of mudrocks: application to Alpine Tethydes shales (Sicily, Italy). *Sediment. Geol.* 213, 27–40.
- Barbera, G., Critelli, S., Mazzoleni, P., 2011. Petrology and geochemistry of cretaceous sedimentary rocks of the Monte Soro Unit (Sicily, Italy): constraints on weathering, diagenesis, and provenance. *J. Geol.* 119 (1), 51–68.
- Bini, C., Sartori, G., Wahsha, M., Fontana, S., 2011. Background levels of trace elements and soil geochemistry at regional level in NE Italy. *J. Geochem. Explor.* 109, 125–133.
- Bonardi, G., De Vivo, B., Giunta, G., Lima, A., Perrone, V., Zuppetta, A., 1982. Mineralizzazioni dell'Arco Calabro-Peloritano. Ipotesi genetiche e quadro evolutivo. *Boll. Soc. Geol. It.* 101, 141–155.
- Cabral Pinto, M., Silva, E., Silva, P., Dinis, F., 2017. Rocha Transfer processes of potentially toxic elements (PTE) from rocks to soils and the origin of PTE in soils: a case study on the island of Santiago (Cape Verde). *J. Geochem. Explor.* 183, 140–151.
- CARG-Cartografia Geologica e Geotematica, 2011. Note illustrative della Carta Geologica D'Italia (1:50.000). Fogli 587, 600.
- Catalano, R., Sulli, A., 2011. Walking along a crustal profile across the Sicily fold and thrust belt. In: Post Conference Field Trip 4, 27–29 October, Palermo, Italy. AAPG Internat. Conf. & Exhibition, pp. 23–26 (October, Milan, Italy).
- Chen, M., Ma, L.Q., Harris, W.G., 1999. Baseline concentrations of 15 trace elements in Florida surface soils. *J. Environ. Qual.* 25, 1173–1181.
- 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 provincial area of Napoli (Italy). *Geochem. Explor. Environ. Anal.* 5, 29–40.
- Cicchella, D., Ambrosino, M., Gramazio, A., Coraggio, F., Assunta, M., Caputi, A., Avagliano, D., Albanese, S., 2022. Using multivariate compositional data analysis (CoDA) and clustering to establish geochemical backgrounds in stream sediments of an onshore oil deposits area. The Agri River basin (Italy) case study. *J. Geochem. Explor.* 238, 107012.
- Crane, J.L., 2021. Distribution and toxic potential of alkylphenols, nonylphenol ethoxylates, and pyrethroids in Minnesota. USA lake sediments. *Sci. Total Environ.* 776, 145974.
- Cressie, N., 1990. The origins of kriging. *Math. Geol.* 239–252.
- D. Lgs, 2006. Legislative Decree 152/2006. Norme in materia ambientale. Allegato 5, Parte IV, Tabella 1. Gazzetta Ufficiale n. 88 del 14 aprile 2006, Supplemento Ordinario n. 96.
- De Lima, M.W., Hamid, S.S., de Souza, E.S., Teixeira, R.A., Palheta, D.d.C., Faial, K.d.C.F., Fernandes, A.R., 2020. Geochemical background concentrations of potentially toxic elements in soils of the Carajás Mineral Province, southeast of the Amazonian Craton. *Environ. Monit. Assess.* 192, 649.
- De Vivo, B., 1982. Mineral resources of the Calabrian-Peloritani Arc: genetic aspect in the evolution of the Arc. *Earth Evolution Sciences* 3, 187–196.
- De Vivo, B., Lima, A., Catalano, G., Chersicla, A., 1993. Detailed geochemical survey in the Peloritani Arc (northeastern Sicily), Italy: evidence of gold anomalies. *J. Geochem. Explor.* 46, 309–324.
- Deutsch, C., 1989. DECLUS: A fortran 77 program for determining optimum spatial declustering weights. *Comput. Geosci.* 325–332.
- Deutsch, C.V., Journel, A.G., 1998. GSLIB. Geostatistical Software Library and User's Guide. Oxford University Press.
- Dogliani, C., Gueguen, E., Harabaglia, P., Mongelli, F., 1999. On the origin of westdirected subduction zones and applications to the western Mediterranean. In: Durand, B., Jolivet, L., Horváth, F., Séranne, M. (Eds.), *The Mediterranean Basins: Tertiary Extension Within the Alpine Orogen*, Geological Society, London, Special Publications, vol. 156, pp. 541–561.
- Dongarrà, G., Ferla, P., 1982. Le argille di portella colla e del flysch numidico auct. (M. Madonie - Sicilia). *Società Italiana di Mineralogia e Petrografia* 38 (3), 1119–1133.
- Dongarrà, G., Manno, E., Sabatino, G., Varrica, D., 2009. Geochemical characteristics of waters in mineralised area of Peloritani Mountains (Sicily, Italy). *Appl. Geochem.* 29, 900–914.
- Dung, T.T.T., Cappuyns, V., Swennen, R., Phung, N.K., 2013. From geochemical background determination to pollution assessment of heavy metals in sediments and soils. *Rev. Environ. Sci. Biotechnol.* 12, 335–353.
- Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York.
- Fabian, C., Reimann, C., Fabian, K., Birke, M., Baritz, R., Haslinger, E., The GEMAS Project Team, 2014. GEMAS: spatial distribution of the pH of European agricultural and grazing land soil. *Appl. Geochem.* 48, 207–216.
- Facchinelli, A., Sacchi, E., Mallen, L., 2001. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environ. Pollut.* 114, 313–324.
- Ferla, P., Omenetto, P., 2000. Metallogenic evolution of Peloritani Mountains (NE Sicily): a summary. *Mem. Soc. Geol. It.* 55, 293–297.
- Fernandes, A.R., Souza, E.S., Braz, A.M.d.S., Birani, S.M., Alleoni, L.R.F., 2018. Quality reference values and background concentrations of potentially toxic elements in soils from the Eastern Amazon. Brazil. *J. Geochem. Explor.* 190, 453–463.
- Frattini, P., De Vivo, B., Lima, A., Cicchella, D., 2006. Background and baseline values of human health harmful elements and gamma-ray survey in the volcanic soils of Ischia island (Italy). *Geochem. Explor. Environ. Anal.* 6, 325–339.
- Galan, E., Fernandez-Caliani, J.C., Gonzalez, I., Aparicio, P., Romero, A., 2008. Influence of geological setting on geochemical baselines of trace elements in soils. Application to soils of South-West Spain. *J. Geochem. Explor.* 98, 89–106.
- Gatuszka, A., 2007. A review of geochemical background concepts and an example using data from Poland. *Environ. Geol.* 52, 861–870.
- Gatuszka, A., Migaszewski, Z.M., 2011. Geochemical background – an environmental perspective. *Mineralogia* 42 (1), 7–17.
- Giacalone, A., Gianguzza, A., Dongarrà, G., Orecchio, S., Piazzese, D., Sciarrino, S., Varrica, D., 2005. Metal distribution in the organic and inorganic fractions of soil: a case-study on soils from Sicily. *Chem. Spec. & Bioavailability* 17 (3), 83–94.
- Gómez-Hernández, J.J., Srivastava, R.M., 2021. One step at a time: the origins of sequential simulation and beyond. *Math. Geosci.* 53, 193–209. <https://doi.org/10.1007/s11004-021-09926-0>.
- Goovaerts, P., 1997. *Geostatistics for Natural Resources*. Oxford University Press.
- Gruba, P., Socha, J., 2016. Effect of parent material on soil acidity and carbon content in soils under silver fir (*Abies alba* Mill.) stands in Poland. *Catena* 140, 90–95.
- He, Z.L., Yang, X.E., Stoffella, P.J., 2005. Trace elements in agroecosystems and impacts on the environment. *J. Trace Elem. Med. Biol.* 19 (2–3), 125–140.
- XLSTAT (versione 2023.3). Addinsoft, Parigi, Francia. <https://www.xlstat.com/en/>.
- Karn, A., Zou, C., Brooks, S., Fresno-Ramírez, J., Gabler, F., Sun, Q., Rammig, D., Naegele, R., Ledbetter, C., Cadle-Davidson, L., 2021. Discovery of the REN11 locus from *Vitis aestivalis* for stable resistance to grapevine powdery mildew in a family segregating for several unstable and tissue-specific quantitative resistance loci. *Front. Plant Sci.* 12, 733899.
- Kicińska, A., Turek, K., 2017. Establishing geochemical background of elements present in soil and its application in the evaluation of soil pollution based on data collected in the Beskid Sądecki region. *Geoinformatica Pol.* 16, 87–99.
- Koppek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* 15 (3), 259–263.
- Kubier, A., Wilkin, R.T., Pichler, T., 2019. Cadmiun in soils and groundwater: a review. *Appl. Geochem.* 108, 104388.
- Li, Q., Wang, Y., Li, Y., Li, L., Tang, M., Hu, W., Chen, L., Ai, S., 2022. Speciation of heavy metals in soils and their immobilization at micro-scale interfaces among diverse soil components. *Sci. Total Environ.* 825, 153862.
- Liguori, V., Brucculeri, R., 2004. Groundwater management resources in Sicily, Italy. http://balwois.com/balwois/administration/full_paper/fp-546.pdf.
- Manly, B.F.J., 1997. *Randomization, Bootstrap, and Monte Carlo Methods in Biology*, Second edition. Chapman Hall, London.
- Matschullat, J., Ottenstein, R., Reimann, C., 2000. Geochemical background – can we calculate it? *Environ. Geol.* 39 (9).
- Mollo, S., Lanzafame, G., Masotta, M., Iezzi, G., Ferlito, C., Scarlato, P., 2011. Cooling history of a dike as revealed by mineral chemistry: a case study from Mt. Etna volcano. *Chem. Geol.* 288, 39–52.
- Müller, G., 1969. Index of geoaccumulation in sediments of the RhineRiver. *Geoj* 2, 108–118.
- Panagos, P., Van Liedekerke, M., Borrelli, P., Köninger, J., Ballabio, C., Orgiazzi, A., Lugato, E., Liakos, L., Hervas, J., Jones, A., Montanarella, L., 2022. European Soil Data Centre 2.0: soil data and knowledge in support of the EU policies. *Eur. J. Soil Sci.* 73 (6), e13315.
- Peccerillo, A., 2005. Plio-Quaternary Volcanism in Italy. *Petrology, Geochemistry, Geodynamics*.

- Plant, J., Smith, D., Smith, B., Williams, L., 2001. Environmental geochemistry at the global scale. *Appl. Geochem.* 16, 1291–1308.
- Pourret, O., Hursthouse, A., 2019. It's time to replace the term "heavy metals" with "potentially toxic elements" when reporting environmental research. *Int. J. Environ. Res. Public Health* 16 (22), 4446.
- R Core Team, R, 2021. *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available online: <https://www.R-project.org/>. (Accessed 1 January 2022).
- Rambeau, C.M.C., Baize, D., Saby, N., Matera, V., Adatte, T., Föllmi, K.B., 2010. High cadmium concentrations in Jurassic limestone as the cause for elevated cadmium levels in deriving soils: a case study in Lower Burgundy, France. *Environ. Earth Sci.* 61, 1573–1585.
- Ramos-Miras, J.J., Díaz-Fernández, P., SanJosé-Wery, A., Rodríguez-Martín, J.A., Roca, N., Bech, J., Roca-Perez, L., Boluda, R., Gil, C., 2014. Influence of parent material and soil use on arsenic forms in soils: a case study in the Amblés Valley (Castilla-León, Spain). *J. Geochem. Explor.* 147, 260–267.
- Reimann, C., de Caritat, P., 2000. Intrinsic flaws of element enrichment factors (EFs) in environmental geochemistry. *Environ. Sci. Technol.* 34, 5084–5091.
- Reimann, C., 2005. Distinguishing between natural and anthropogenic sources for elements in the environment: regional geochemical surveys versus enrichment factor. *Sci. Total Environ.* 337, 91–107.
- Reimann, C., Garrett, R.G., 2005. Geochemical background – concept and reality. *Sci. Total Environ.* 350, 12–27.
- Reimann, C., Filzmoser, P., Garrett, R.G., 2005. Background and threshold: critical comparison of methods of determination. *Sci. Total Environ.* 346, 1–16.
- Reimann, C., Birke, M., Demetriades, A., Filzmoser, P., O'Connor, P., 2014. *Chemistry of Europe's Agricultural Soils, Part B: General Background Information and Further Analysis of the GEMAS Data Set*, Geologisches Jahrbuch Reihe B, Band B 103. Schweizerbart Science Publishers, Stuttgart, p. 352.
- Reuter, H.I., Lado, L.R., Hengl, T., Montanarella, L., 2008. Continental-scale digital soil mapping using European soil profile data: soil pH. *Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie* 19, 91–102.
- Saccà, C., Saccà, D., Nucera, P., De Fazio, A., 2011. Composition and geochemistry of clay sediments offshore the northeastern Sicilian coast (Southeastern Tyrrhenian Sea, Italy). *Estuar. Coast. Shelf Sci.* 92, 564–572.
- Salminen, R., Gregorauskiene, V., 2000. Considerations regarding the definition of a geochemical baseline of elements in the surficial materials in areas differing in basic geology. *Appl. Geochem.* 15, 647–653.
- Salminen, R., Tarvainen, T., 1997. The problem of defining geochemical baselines. A case study of selected elements and geological materials in Finland. *J. Geochem. Explor.* 60 (1), 91–98.
- Santos-Francés, F., Martínez-Graña, A., Alonso Rojo, P., García, Sánchez A., 2017. Geochemical background and baseline values determination and spatial distribution of heavy metal pollution in soils of the Andes Mountain Range (Cajamarca-Huancavelica, Peru). *Int. J. Environ. Res. Public Health* 14 (8), 859 (31).
- Sappa, G., Barbieri, M., Andrei, F., 2020. Assessment of trace elements natural enrichment in topsoil by some Italian case studies. *SN Appl. Sci.* 2, 1409.
- Sierra, C., Ordóñez, O., Saavedre, A., Gallego, J.R., 2015. Element enrichment factor calculation using grain-size distribution and functional data regression. *Chemosphere* 119, 1192–1199.
- Singh, A., Maichle, R., 2015. *ProUCL Version 5.1 User Guide*. U.S. Environmental Protection Agency, Washington, DC. https://www.epa.gov/sites/production/files/201605/documents/proucl_5.1_user-guide.pdf (EPA/600/R-07/041).
- SNPA – Sistema nazionale per la protezione ambientale. Linea guida per la determinazione valori di fondo per i suoli e per le acque sotterranee. https://www.isprambiente.gov.it/files2018/publicazioni/manuali-linee-guida/MLG_174_18.pdf (file PDF: last consultation 19 July 2024).
- Souza, E.S., Texeira, R.A., da Costa, H.S.C., Oliveira, F.J., Melo, L.C.A., Faial, K.d.C.F., Fernandes, A.R., 2017. Assessment of risk to human health from simultaneous exposure to multiple contaminants in an artisanal gold mine in Serra Pelada, Pará, Brazil. *Sci. Total Environ.* 576, 683–695.
- Suresh, G., Sutharsan, P., Ramasamy, V., Venkatachalapathy, R., 2012. Assessment of spatial distribution and potential ecological risk of the heavy metals in relation to granulometric contents of Veranam lake sediments. *India. Ecotoxicol. Environ. Saf.* 84, 117–124.
- USDA - United States Department of Agriculture. *Soil Survey Manual*. 2017. Handbook No. 18. <https://www.nrcs.usda.gov/sites/default/files/2022-09/The-Soil-Survey-Manual.pdf> (file PDF: last consultation 10 July 2024).
- Varrica, D., Lo Medico, F., Alaimo, M.G., 2022. Air quality assessment by the determination of trace elements in lichens (*Xanthoria calcicola*) in an industrial area (Sicily, Italy). *Int. J. Environ. Res. Public Health* 19, 9746.
- Wei, C., Wen, H., 2012. Geochemical baselines of heavy metals in the sediments of two large freshwater lakes in China: implications for contamination character and history. *Environ. Geochem. Health* 34, 737–748.
- Xu, D., Shen, Z., Dou, C., Dou, Z., Li, Y., Gao, Y., Sun, Q., 2022. Effects of soil properties on heavy metal bioavailability and accumulation in crop grains under different farmland use patterns. *Sci. Rep.* 12, 9211.
- Yongming, H., Peixuan, D., Junji, C., Posmentier, E.S., 2006. Multivariate analysis of heavy metal contamination in urban dusts of Xi'an, Central China. *Sci. Total Environ.* 355 (1–3), 176–186.
- Zhang, Y.Y., Wu, W., Liu, H., 2019. Factors affecting variations of soil pH in different horizons in hilly regions. *PLoS One* 14 (6), e0218563.
- Zoller, W.H., Gladney, E.S., Duce, R.A., 1974. Atmospheric concentrations and sources of trace metals at the South Pole. *Science* 183, 198–200.