

Article

Regional Baseline Estimation in Campania, Southern Italy: Incorporating Spatial Autocorrelation via Hotspot Analysis

Antonio Iannone ^{1,*} , Salvatore Dominech ^{1,*} , Chaosheng Zhang ² and Stefano Albanese ¹ 

¹ Department of Earth Sciences, Environment and Resources (DiSTAR), University of Naples Federico II, 80126 Naples, Italy; stefano.albanese@unina.it

² International Network for Environment and Health (INEH), School of Geography, Archaeology & Irish Studies, University of Galway, H91 CF50 Galway, Ireland; chaosheng.zhang@universityofgalway.ie

* Correspondence: antonio.iannone2@unina.it (A.I.); salvatore.dominech@unina.it (S.D.)

Abstract

This study applies a hotspot-based spatial statistical approach to investigate the spatial distribution of chemical elements and to improve regional geochemical baseline estimation in topsoils affected by widespread anthropogenic influence. Specifically, this study applied the Getis–Ord Gi* Hotspot analysis on over 7000 topsoil samples from the Campania region (southern Italy), focusing on 21 variables. The analysis revealed statistically significant clusters of high and low concentrations, closely aligned with regional geological features. Elevated levels of As, Ba, Be, Bi, Cu, Sr, Th, Tl, U, and V were mainly observed in soils developed on volcanoclastic deposits, whereas Co, Cr, Ni, and Mn were more common in soils on siliciclastic units. Cd, Hg, Pb, Sb, Sn, and Zn exhibited clustered anomalies in major urban and industrial areas, indicating anthropogenic sources. For these elements, baseline values were estimated. Traditional statistical methods, which primarily rely on data distribution, often overlook spatial autocorrelation, leading to biased thresholds, particularly in areas with widespread contamination. The hotspot-based approach addresses this limitation by excluding hotspot clusters from the calculation of the 95% Upper Tolerance Limit (UTL95-95), thereby providing baseline thresholds uninfluenced by human activity. Comparison with other data-driven methods showed consistent trends across lithologies, although the hotspot-based approach tended to yield slightly lower thresholds, reflecting its responsiveness to spatial patterns.

Keywords: soil geochemistry; spatial clustering; geogenic controls; anthropogenic impacts; environmental thresholds



Academic Editor: Joaquim Esteves Da Silva

Received: 16 December 2025

Revised: 4 February 2026

Accepted: 7 February 2026

Published: 11 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

In recent decades, environmental geochemistry has played a crucial role in understanding and managing both natural ecosystems and mineral resources. Its methodologies significantly reduce the environmental impacts of resource extraction, soil erosion, and water pollution [1]. Additionally, this discipline offers several tools for addressing environmental tasks, including pollution remediation and climate change mitigation, by improving the understanding of how chemicals behave, move, and transform in natural systems. Among all chemicals, Potentially Toxic Elements (PTEs) are of particular concern because of their persistence in the environment and well-known adverse effects on both ecosystems and human health [2] when above specific thresholds. In this framework, soils represent reservoirs of PTEs and compounds derived from both geogenic and anthropogenic sources. The first are usually displayed in regional-scale geochemical patterns and include

weathering and soil formation processes. Anthropogenic factors, which are generally more localized, include human activities such as traffic, waste disposal, agriculture, and industrialization [3]. A key challenge in environmental geochemistry is precisely differentiating between natural concentrations of PTEs derived from the underlying geological setting and anomalies resulting from human activities. This is especially important in environmental monitoring, where detecting contamination enables understanding of what contributes to the natural background/baseline, and in resource exploration, where such information may represent hidden mineralization [4,5].

A second critical objective in topsoil geochemistry is the estimation of natural background or baseline concentrations, which serve as a reference for identifying anomalies or contaminated sites [6,7]. Specifically, geochemical background refers to the natural abundance of an element in a given geologic matrix (e.g., soil, sediment, or rock) within a specific area or dataset, reflecting concentrations controlled by geological and pedogenetic processes. A geochemical baseline is a statistical representation of element concentrations in a region that may be affected by diffuse anthropogenic inputs. Unlike the geochemical background, the baseline is an operational reference level intended for environmental assessment and anomaly detection in impacted areas [8]. This is a complex task because data on anthropogenic inputs tend to skew the dataset's geochemical distribution, increasing both central tendency and variance, particularly in the case of diffuse contamination [9]. This displacement may result in poorly defined thresholds, reducing anomaly detection sensitivity and potentially underestimating environmental risk [10].

Numerous statistical approaches have been proposed to overcome perturbations caused by human pressure in background/baseline assessment. Standard techniques include the Upper Tolerance Limit (UTL) at 95% confidence [11], the Mean $\pm$ 2 Standard Deviations (2σ method) [12], Iterative 2σ filtering [9], Cumulative Frequency Curves (CFC) [13], boxplots [13], and percentile-based thresholds [9]. More advanced techniques, such as Concentration–Area (C–A) analysis [14] and Spectrum–Area (S–A) analysis [15,16], have also been used. In particular, the use of UTL95 in environmental regulations is most appropriate when natural background or baseline values are determined for contaminants in soils and groundwater. In fact, in Italy, the National System for Environmental Protection (Sistema Nazionale per la Protezione dell'Ambiente—SNPA) recommends adopting the UTL95 as a statistical upper limit, representing the concentration below which the 95% of the natural population is expected to fall, with a 95% level of confidence. However, all these methods, which can be categorized as data-driven, rely solely on the dataset's intrinsic properties. Although they are effective in homogeneous or uncontaminated regions, their reliability decreases when the dataset includes numerous samples influenced by anthropogenic activities [17]. In such cases, background/baseline values can be biased upward, leading to thresholds that fail to distinguish anomalous sites in the case of diffuse contamination. Conversely, overly restrictive filtering can remove part of the geological control, leading to an underestimation of the threshold levels. Another critical limitation of these approaches is their lack of spatial consideration, which assumes that all observations are independent and equally distributed, thereby ignoring spatial autocorrelation, a key feature in geochemical data. The spatial distribution of chemical concentrations, especially in the presence of widespread contamination, is not random but tends to produce clustered anomalies with distinct spatial structures. Ignoring these spatial patterns may reduce the effectiveness of traditional statistical methods and lead to threshold values that include historically contaminated areas [18]. Despite growing attention to the spatial analysis in environmental studies, autocorrelation-based techniques have rarely been applied to the specific task of geochemical background/baseline estimation in regions where contamination is not isolated but occurs as spatially coherent patterns, such as urbanized

or industrial areas. Integrating spatial autocorrelation into the background assessment process could significantly enhance the accuracy of its determination in areas where contaminant concentrations generate statistically significant clusters, thereby allowing more realistic thresholds and improving the spatial identification of anomalous geochemical signatures. In complex and heterogeneous regions, a high sampling density and sample size could therefore represent the most effective approach to ensure adequate control over the robustness of background/baseline estimation.

Given the above considerations and building upon a previous work [18] that applied PCA to a regional dataset of topsoil samples of the Campania region (southern Italy), the present study aims to perform a complementary follow-up by using the spatial statistical tool based on Hotspot Analysis (Getis–Ord Gi*) [19]. This technique identifies statistically significant clusters of high and low values across a given area, enabling a more refined view and interpretation of the spatial pattern [20]. In addition to PCA, which provides a global perspective based on element associations, Hotspot analysis captures spatial heterogeneity, highlights specific areas of concern, and supports the exploration of geochemical processes driven by hypotheses.

The Campania region in Southern Italy has been selected as a case study to examine the spatial patterns of contamination processes. We applied hotspot analysis as a filtering tool to the geochemical elements primarily associated with anthropogenic activity. Consequently, we estimated baseline values for the selected elements (PTEs) using the Upper Tolerance Limit (UTL) approach at 95% confidence (UTL95-95; [11]). Finally, we compared the resulting baseline values with those reported in the literature to assess the significance of the dataset's spatial structure and its impact on baseline estimation. This study presents the first integration of local spatial statistics into the preprocessing of topsoil data before regional baseline calculation in the Italian context, explicitly accounting for spatial autocorrelation.

2. Materials and Methods

2.1. Study Area

The Campania Region (southern Italy—Figure 1) spans approximately 13,600 km² and ranks as the third-most populous region in Italy, with a population of roughly 5.6 million. The entire territory is traversed by the NW-SE-oriented Apennine Mountain chain, which formed during the Cretaceous–Tertiary phase of the Alpine orogeny. The Apennine units are mainly characterized by Mesozoic carbonate successions and Tertiary siliciclastic deposits, reflecting deposition in a variety of tectono-sedimentary settings. The highest elevations, reaching up to 1800–2000 m, are dominated by carbonate lithologies and exhibit steep slopes, often mantled by colluvial and debris-flow deposits. These mountainous areas are surrounded by hilly terrains of moderate elevation, where siliciclastic rocks such as sandstones, siltstones, claystones, and conglomerates are prevalent (Figure 1a). During the Quaternary, magmatic activity began with the formation of the Roccamonfina volcano (700 ka), followed by the volcanic systems of Ischia Island (150 ka), the Phlegraean Fields (60 ka), and the Somma–Vesuvius complex (25 ka). The oldest volcanic units are found in the Roccamonfina area, where ultrapotassic to potassic lavas and pyroclastic products were emplaced between 0.6 and 0.1 Ma [21,22]. The most voluminous pyroclastic accumulations are associated with the explosive activity of the Phlegraean Fields and Mt. Somma–Vesuvius, particularly the major ignimbrite-forming eruption that occurred approximately 39 ka [23–25].

This geological complexity is clearly reflected in the regional geochemical pattern of topsoils and represents a significant geogenic source of elemental variability. Specifically, volcanic lithologies are responsible for natural enrichments of elements such as Be, Tl, Th, and U. In contrast, siliciclastic units are commonly associated with higher concentrations

of Mn, Ni, and Co. The study area is also subject to significant anthropogenic pressure, mainly localized around the main urban centers. The primary sources of contamination include historical and ongoing industrial activities (e.g., metal processing, chemical and textile industries), waste disposal and illegal dumping, intense road traffic, and intensive agriculture with high fertilizer and pesticide use [26], contributing to the accumulation of PTEs in topsoil over the decades.

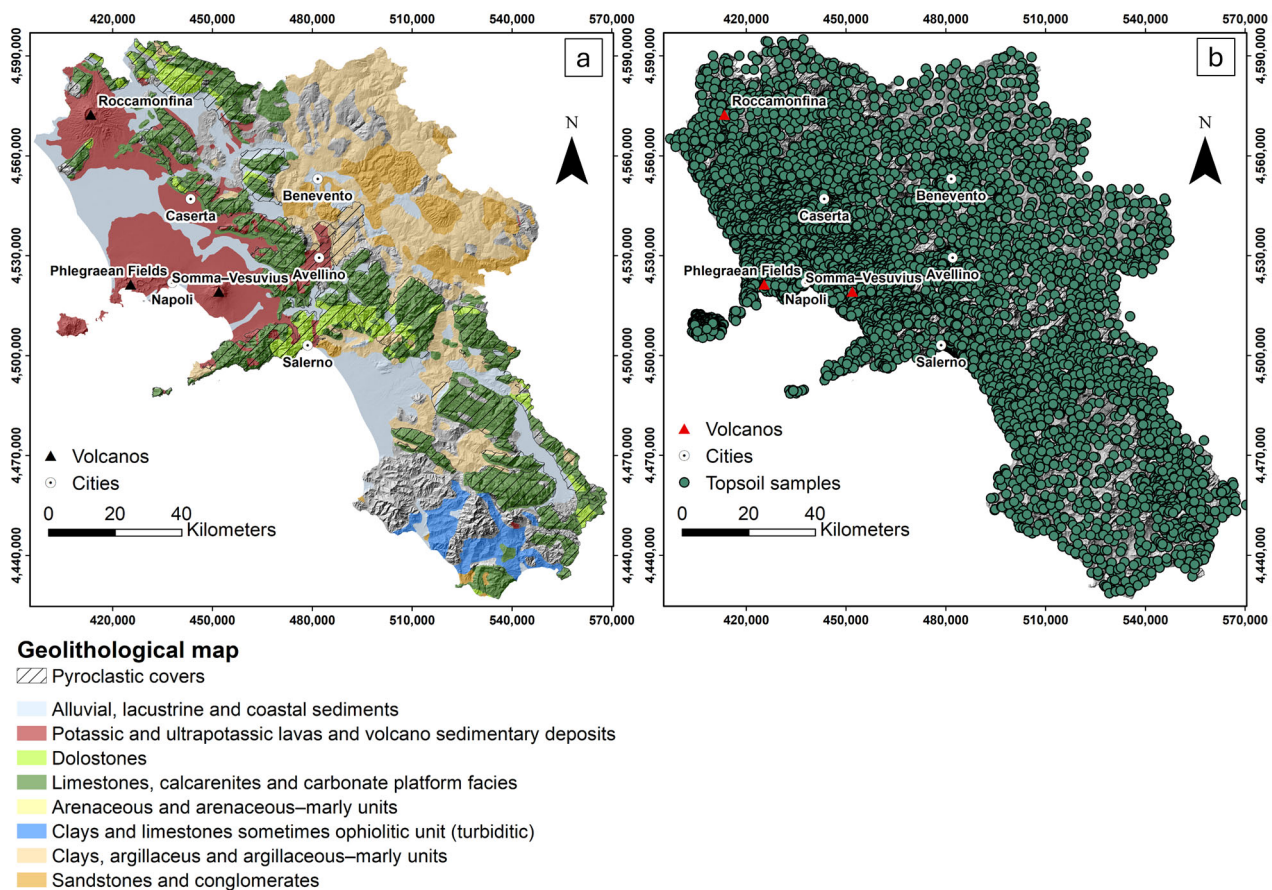


Figure 1. Study area characterization with (a) geolithological map showing the distribution of volcanic, alluvial, siliciclastic, and carbonatic units across the regional territory (modified after Iannone et al. [18] and reference therein); (b) topsoil samples distribution map. The coordinate reference system used for all the maps is WGS84 UTM Zone 33N.

2.2. Sampling and Chemical Analysis

Soil sampling for this study followed standardized protocols recommended by major European geochemical mapping initiatives, specifically those established by the Forum of European Geological Surveys (FOREGS, now EuroGeoSurveys) [27], as well as the GEMAS (Geochemical Mapping of Agricultural and Grazing Land Soil) project [20]. The adopted methodology involved a grid-based sampling in which the territory was divided into cells of variable size, depending on the degree of environmental pressure. Specifically, 4×4 km grid cells were used in forested areas characterized by low environmental pressure, 2×2 km predominantly in agricultural regions with moderate pressure, and 1×1 km in urban areas with high environmental pressure. Within each cell, five subsamples were collected (one from each corner and one from the center) and subsequently homogenized into a single composite sample.

In total, 7300 topsoil samples (collected at 5–15 cm depth) were collected across the study area (Figure 1b). The dataset comprises two main subsets acquired through successive geochemical prospecting campaigns ([28] and references therein), which were

later integrated into a single regional framework. The sampling campaigns were carried out from 2015 to 2017. The samples were placed in polyethylene bags and transported to the laboratory for analysis. Upon arrival, samples were air-dried under infrared lamps, with thermocouples used to ensure the samples remained below 35 °C. They were then sieved to separate the <2 mm fraction.

These prepared samples were analyzed for a suite of 52 elements, including major, minor, and trace elements (Ag, Al, As, Au, B, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cs, Cu, Fe, Ga, Ge, Hf, Hg, In, K, La, Li, Mg, Mn, Mo, Na, Nb, Ni, P, Pb, Pd, Pt, Rb, Re, S, Sb, Sc, Se, Sn, Sr, Te, Th, Ti, Tl, U, V, W, Y, Zn, Zr), using both Inductively Coupled Plasma Mass Spectrometry (ICP-MS) and Inductively Coupled Plasma Emission Spectrometry (ICP-OES). Specifically, a 15 g split of the pulp was digested in 45 mL of aqua regia mixture at 90 °C for one hour (pseudototal digestion). The resulting solution was then brought to a final volume of 300 mL with 5% HCl. All analytical procedures were carried out by Bureau Veritas Analytical Laboratories (formerly Acme Labs) in Vancouver, Canada. Throughout the sampling project, the analytical protocol was progressively expanded, with new elements incorporated into the detection suite as laboratory capabilities evolved. Consequently, the number of available measurements varies across elements, as some were introduced only during later phases of the study.

To verify data quality and analytical reliability, the laboratory employed standard quality control procedures, including the use of reagent blanks, certified reference materials, and replicate (duplicate) analyses. Analytical precision was quantified by calculating the Relative Percent Difference (%RPD) between duplicate samples, while accuracy was assessed by comparing measured values with certified reference standards. When elemental concentrations fell below the method detection limit (DL), values were set to half the DL, an approach commonly adopted to minimize bias without assuming the analyte is absent [12]. Details on analytical reproducibility and accuracy are reported in the Results section.

2.3. Data Preparation and Hotspot Analysis

A subset of 21 elements (As, Ba, Be, Bi, Cd, Co, Cr, Cu, Hg, Ni, Mn, Mo, Pb, Sb, Sn, Sr, Th, Tl, U, V, and Zn) out of the original 52 was used. The selected elements included well-known PTEs and other elements, such as Be, Bi, Sr, and Th, recognized as markers in characterizing the geochemical signal of our region of interest [28,29]. This selection also reflected continuity with the previous work by [18], ensuring consistency and comparability between the two studies.

A Normal Score Transformation (NST) was applied to the dataset as a preprocessing step to address non-normality and enable hotspot analysis. This normalization technique was employed to reduce the influence of extreme values and to approximate the original data distribution to a standard normal distribution (mean = 0; standard deviation = 1) by ranking the data values from lowest to highest. Compared to other normalization approaches, its non-linear score transformation offers the advantage of improving any datasets for variance-based and spatial statistical tools such as the Getis–Ord G_i^* [20]. A Kolmogorov–Smirnov (K-S) test was also conducted on the NST transformed data to evaluate the effectiveness of the transformation. The data transformation and K-S test were performed using IBM SPSS Statistics 25, while Microsoft Excel was utilized for descriptive statistical analysis.

The hot-spot analysis was applied to the normalized dataset for the 21 selected elements. This spatial statistical technique can uncover patterns of spatial clustering based on sample distances, identifying areas characterized by statistically significant concentrations of high (hotspot) and low (coldspot) values. The analysis is based on the Getis–Ord G_i^* statistic, a local measure of spatial association that quantifies the degree of clustering for

each feature relative to its neighbors [19]. The Gi* statistic produces z-scores and *p*-values, enabling the identification of features surrounded by values that are similarly high or low. High z-scores related to low *p*-values indicate statistically significant hotspots, while low (negative) z-scores related to low *p*-values identify the presence of coldspots. The function of the Getis–Ord Gi* statistics is defined as follows [30]:

$$G_i^* = \frac{\sum_{j=1}^n \omega_{ij} x_j - \bar{X} \sum_{j=1}^n \omega_{ij}}{SD \sqrt{\frac{[n \sum_{j=1}^n \omega_{ij}^2 - (\sum_{j=1}^n \omega_{ij})^2]}{n-1}}} \quad (1)$$

where *i* is the centre of a defined neighborhood, *x_j* is the value concentration at the sample location *j*, *ω_{i,j}* is the spatial weight between locations *i* and *j*, and *n* is the total number of samples. The mean of the dataset $\bar{X}$ is calculated with the following equation:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

The standard deviation *SD* of the dataset is calculated through the following equation:

$$SD = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3)$$

Spatial relationships were conceptualized using an inverse-distance weighting scheme based on Euclidean distance, whereby observations closer together have a stronger influence than those farther apart. The neighborhood search threshold, automatically determined by ArcMap (version 10.8), was set at 4390 m to ensure that each feature had at least one neighbor without imposing a fixed number of neighbors, which could introduce artificial spatial connections in sparsely sampled areas. The default row-standardized weighting matrix mitigated potential bias arising from varying numbers of neighboring features, given the dataset's heterogeneous sampling density, with densely sampled urban areas and sparsely sampled mountainous and hilly zones. The hotspot analysis and mapping were performed using the Hotspot Analysis (Getis–Ord Gi*) tool in ArcMap (version 10.8).

2.4. Regional Geochemical Baseline Estimation

Six PTEs (Zn, Sb, Hg, Pb, Sn, and Cd) out of the 21 selected elements were selected for further investigation, as they were associated with urban and industrial sources by Iannone et al. [18]. For each of these six variables, the Getis–Ord Gi* results were used to identify and exclude statistically significant hotspots, defined as samples with Gi-bin values of 1, 2, or 3, corresponding to 90%, 95%, and 99% confidence levels, respectively. These confidence intervals correspond to *p*-values (0.1, 0.05, and 0.01, respectively) that indicate the probability that the observed spatial cluster occurred by chance, rather than as a result of an underlying spatial process. Samples interpreted as potentially impacted by site-specific contamination or locally enriched by natural processes were excluded to avoid bias in the estimation of the regional geochemical baseline. The remaining dataset, representing values very close to the average regional geochemical trend, was further subdivided into four subsets based on the main lithological categories of the study area: volcanic, carbonate, siliciclastic, and alluvial units. This subdivision was chosen because lithology exerts primary control over soil geochemical signatures, and significant variations in element associations can be effectively observed only within these major lithological groups. Each lithological subset was then subjected to a Box–Cox transformation. This widely used normalization method enables the identification of an optimal power transformation that transforms the raw data distribution into an approximate normal distribution [31]. Since

the NST is a non-linear transformation based on ranking and does not allow the back-transformation to raw concentrations, the Box–Cox method has been preferred for this purpose. Further, the 95% Upper Tolerance Limit (UTL95-95) was calculated to estimate baseline concentrations. This statistical threshold provides a conservative upper boundary below which 95% of the natural population is expected to fall, with 95% confidence, offering a reliable upper threshold for baseline estimation in areas with diffuse contamination. Finally, the calculated UTL95-95 values were back-transformed to raw concentrations, yielding baseline estimates for each of the six selected PTEs across the four major lithologies. The Box–Cox transformation and the UTL95-95 calculation were performed in R (version 2022.07.2, Build 576) and Pro UCL 5.2.0.0, respectively.

3. Results

Principal statistical indices for the 21 elements are summarized in Table 1. The wide range between the minimum and maximum values, as well as the high coefficient of variation (CV), suggests considerable variability throughout the dataset. Positive skewness and high kurtosis values highlight the presence of extreme concentration and non-normal distributions. For most samples, accuracy and precision values were within 5% and 10%, respectively. In a limited number of cases, higher values were observed; nevertheless, they did not exceed 20%, the commonly accepted threshold for environmental geochemical quality control.

Table 1. Summary of descriptive and performance statistics for each analyzed element. Reported parameters include the number of samples, minimum value, 25th percentile (P25), median, mean, 75th percentile (P75), 95th percentile (P95), maximum value, standard deviation (SD), coefficient of variation (CV%), median absolute deviation (MAD), skewness, kurtosis, analytical accuracy, and relative percentage difference (RPD). Element concentrations are expressed in milligrams per kilogram (mg/kg), except for mercury (Hg), which is given in micrograms per kilogram (µg/kg).

Element	Samples	Min	25	Median	Mean	P75	P95	Max	SD	CV%	MAD	Skewness	Kurtosis	Accuracy (%)	RPD (%)
As	7300	0.60	7.90	11.8	12.4	15.2	21.9	930	13.0	105	3.70	48.9	3387	2.60	2.70
Ba	7300	8.30	194	337	368	505	762	2953	220	59.7	152	1.16	5.19	4.30	5.60
Be	5864	0.10	2.40	4.70	4.56	6.30	8.48	17.9	2.42	53.1	1.80	0.33	−0.04	6.90	11.1
Bi	6963	0.03	0.32	0.42	0.47	0.53	0.85	11.8	0.35	75.4	0.11	13.0	310	8.50	16.2
Cd	7152	<0.01	0.20	0.31	0.45	0.51	1.29	11.1	0.52	114	0.13	6.74	86.4	5.20	6.90
Co	7300	0.50	8.30	11.5	11.8	14.6	20.3	88.0	5.36	45.5	3.20	1.84	13.7	5.50	8.90
Cr	7300	<0.50	10.8	16.9	21.7	27.5	48.3	808	24.1	111	7.40	13.4	312	5.20	4.60
Cu	7300	2.51	32.9	52.9	93.5	111	272	2394	126	134	26.8	6.04	59.7	4.80	8.40
Hg	6915	<5.00	29.0	45.0	82.3	77.0	269	6775	161	196	20.0	16.8	532	9.00	17.7
Mn	7300	51.0	677	834	945	1082	1761	7975	509	53.9	188	4.17	34.1	2.70	3.80
Mo	7300	0.06	0.74	1.10	1.38	1.67	2.82	62.1	1.52	110	0.44	14.3	413	6.00	4.40
Ni	7300	0.40	12.0	16.1	19.9	25.3	45.0	156	13.3	67.2	5.60	1.85	5.82	3.50	3.70
Pb	7300	3.12	29.7	47.1	59.9	66.6	140	2052	70.8	118	18.3	9.20	152	4.80	3.30
Sb	7065	0.01	0.35	0.52	0.79	0.79	2.07	42.8	1.42	179	0.20	14.5	335	14.7	3.50
Sn	5864	0.20	1.70	2.90	3.54	4.10	7.90	126	4.48	127	1.20	13.6	277	5.70	4.70
Sr	7300	4.60	80.0	141	166	223	382	1371	115	69.3	69.3	1.49	5.80	6.90	5.10
Th	7300	0.30	7.10	12.1	12.6	16.4	25.0	64.3	7.16	56.8	4.70	1.24	3.58	6.20	5.60
Tl	7113	0.05	0.71	1.36	1.38	1.95	2.60	69.0	1.14	82.8	0.62	30.0	1753	3.60	3.50
U	7134	<0.1	1.30	3.00	3.35	4.80	7.50	43.2	2.40	71.5	1.70	1.88	14.7	6.90	2.20
V	7300	4.00	45.0	64.0	76.6	89.0	118	234	29.5	43.6	22.0	0.48	−0.11	4.60	10.0
Zn	7300	3.90	68.0	85.9	104	111	210	3211	91.1	87.9	20.4	11.6	263	3.60	4.70

The results of the Kolmogorov–Smirnov (K-S) test (p -value < 0.05), combined with the pronounced positive skewness and kurtosis, confirm the non-normality of the raw dataset. However, the application of the NST yielded a symmetrical distribution (Figure 2), as suggested by the K-S test (p -value > 0.05).

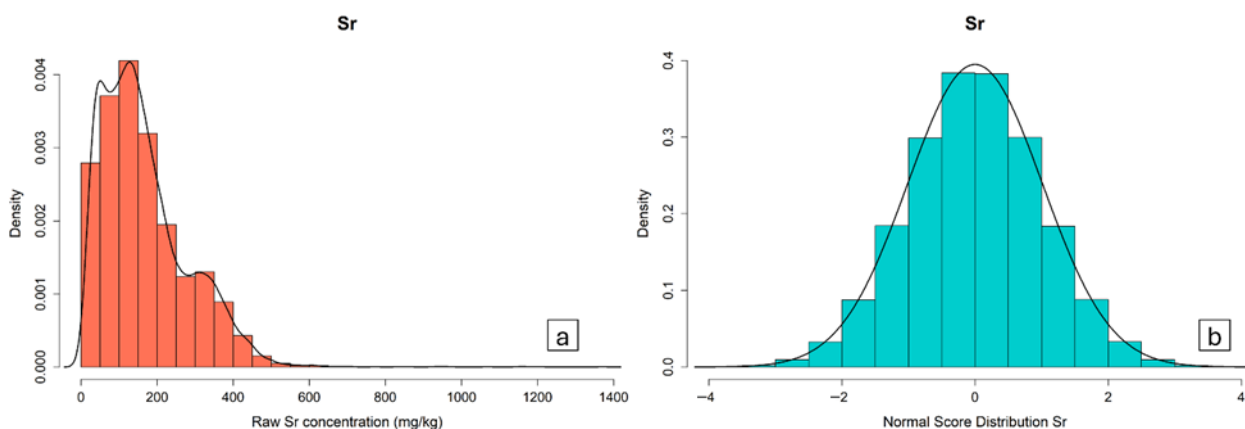


Figure 2. Example using Sr concentration data: (a) histogram of raw Sr concentrations (mg/kg) with superimposed probability density function, highlighting the pronounced positive skewness typical of many geochemical datasets; (b) histogram of the same data after normal score transformation, with superimposed normal probability density function, approximating a standard normal distribution.

The spatial distribution patterns of element concentrations across the study area reveal the influence of both geogenic and anthropogenic factors. Elements such as As, Bi, Be, Mo, Th, and U present relatively higher concentration values in volcanic zones and carbonate massifs covered by pyroclastic materials. As and Bi reach elevated levels in areas affected by Campanian Ignimbrite deposits, while Mo and Th exhibit higher values in the southern sectors and the Roccamonfina area, respectively. Similarly, Co, Ni, Mn, and Cr display medium to high concentrations across the southern and northeastern sector of the region, with Mn showing a decreasing gradient toward the coastal plain. The volcanic area of the Somma–Vesuvius volcano is associated with high values of Ba, Sr, and Cu; in contrast, elements such as Pb, Sb, Sn, Zn, Hg, and Cd demonstrate spatial associations with urbanized and industrial zones. High Pb and Zn values are observed across all major urban centers, whereas elevated concentrations of Hg and Cd are primarily detected in the cities of Naples and Salerno. Thallium and V are distributed in volcanic areas and carbonate uplands, with concentration values decreasing toward the peripheral sectors.

3.1. Hotspot Analysis Getis–Ord G_i^*

The Getis–Ord G_i^* statistics identified statistically significant clusters of hotspots and coldspots for each element, highlighting a combination of natural and anthropogenic drivers behind the distribution of elemental hotspots (Figures 3–5). Hotspots of Be, Bi, Th, Tl, and U are mainly located in the northern sector but also widespread in the southern and central parts, presenting a spatial correspondence with volcanic terrains of Roccamonfina and Somma–Vesuvius areas, as well as carbonate massifs of the Sorrento Peninsula and the Apennine chain, mainly covered by pyroclastic deposits. Coldspots are located across the northeastern and southern sectors of the region, corresponding with siliciclastic units. A similar clustering pattern is observed for As, which also exhibits hotspots in the Phlegraean Fields area. Furthermore, coldspots of As are not limited to the northeastern and southern sectors but are also located around the Somma–Vesuvius area. Several elements, such as Hg, Mo, Pb, Sb, Sn, and Zn, exhibit similar clustering patterns, with the location of hotspots surrounding the major urban centers, especially in the central and coastal zones, and coldspots mainly located in correspondence with the siliciclastic units in the northeast-

ern part of the territory. In addition, Cd shows hotspots located along the northern and central-southern sectors, in correspondence with carbonates mantled by pyroclastic covers. Coldspots in this group of elements characterized the volcanic areas and the northern and southwestern sectors, where siliciclastic units are present. Elements such as Co, Cr, Ni, and Mn exhibit broader, more complex hotspot distributions, with significant clusters in both the northeast and the south and show good spatial correspondence with the siliciclastic units. Additionally, Cr shows a clustering pattern in the area between Avellino, Napoli, and Salerno. Coldspots are primarily located in volcanic areas, particularly the Phlegraean Fields. Cu and Ba show hotspots clustering around the Somma–Vesuvius volcano, extending also across the carbonatic reliefs of Avellino. The coldspots are widespread locally across the remaining territory. Strontium and V exhibit a similar clustering pattern, with hotspots also found locally in the northern and eastern sectors.

3.2. Regional Geochemical Baselines

Based on hotspot filtering, baseline values were estimated for the six elements potentially associated with anthropogenic sources: Cd, Hg, Pb, Sb, Sn, and Zn. Specifically, 364 samples (5%) were removed for Cd, 400 (6.5%) for Hg, 417 (5.7%) for Pb, 383 (5.4%) for Sb, 328 (5.6%) for Sn, and 392 (5.3%) for Zn. The baseline calculation was performed through UTL95–95 on the remaining dataset for each major lithological unit. The resulting baseline values were compared with those calculated by De Vivo et al. [28] and Pacifico et al. [32] on the same starting dataset (Table 2).

Table 2. Baseline values for Cd, Hg, Pb, Sb, Sn, and Zn determined using the UTL95% with 95% coverage (UTL95–95) approach for the four major lithotypes of the Campania region (this study), compared with values reported by De Vivo et al. [28] obtained using the data-driven Spectrum–Area (S–A) analysis and values reported by Pacifico et al. [32] derived using the geochemical gene concept. Concentrations are expressed in mg/kg for Cd, Pb, Sb, Sn, and Zn, and in µg/kg for Hg.

	Alluvial			Carbonatic			Volcanic			Siliciclastic		
	This Study	Pacifico et al. [32]	De Vivo et al. [28]	This Study	Pacifico et al. [32]	De Vivo et al. [28]	This Study	Pacifico et al. [32]	De Vivo et al. [28]	This Study	Pacifico et al. [32]	De Vivo et al. [28]
Cd (mg/kg)	0.80	-	0.80	1.20	-	1.80	1.20	-	0.70	0.70	-	0.90
Hg (µg/kg)	190	353	229	136	243	104	184	337	289	98.9	89.8	74.3
Pb (mg/kg)	98.6	150	92.4	95.2	81	79.2	111	142	156	78.6	69.1	72.2
Sb (mg/kg)	1.30	2.27	1.20	1.30	4.25	1.20	1.50	2	2.00	1.00	1.1	1.20
Sn (mg/kg)	6.30	-	5.90	5.50	-	5.90	6.40	-	7.00	4.20	-	5.00
Zn (mg/kg)	168	234	138	155	133	132	165	238	185	129	131	117

For Cd, the highest baseline concentrations were determined for volcanic (1.20 mg/kg) and carbonatic (1.20 mg/kg) units, while the lowest value was observed in siliciclastic ones (0.70 mg/kg). Similarly, Hg shows higher values in alluvial (190 µg/kg) and volcanic (184 µg/kg) lithologies. Lead showed the highest value in volcanic soils (111 mg/kg), followed by alluvial soils (98.6 mg/kg) and carbonate-pyroclastic soils (95.2 mg/kg), with the lowest in siliciclastic soils (78.6 mg/kg). Antimony has relatively homogeneous values across lithologies, with slightly higher concentrations in volcanic soils (1.50 mg/kg) and slightly lower in siliciclastic soils (1 mg/kg). For Sn, the highest baseline values were observed in volcanic (6.40 mg/kg) and alluvial (6.30 mg/kg) soils, whereas siliciclastic soils showed the lowest values (4.20 mg/kg). Finally, Zn showed the highest baseline values in alluvial (168 mg/kg) and volcanic (165 mg/kg) soils, with the lowest in siliciclastic soils (129 mg/kg).

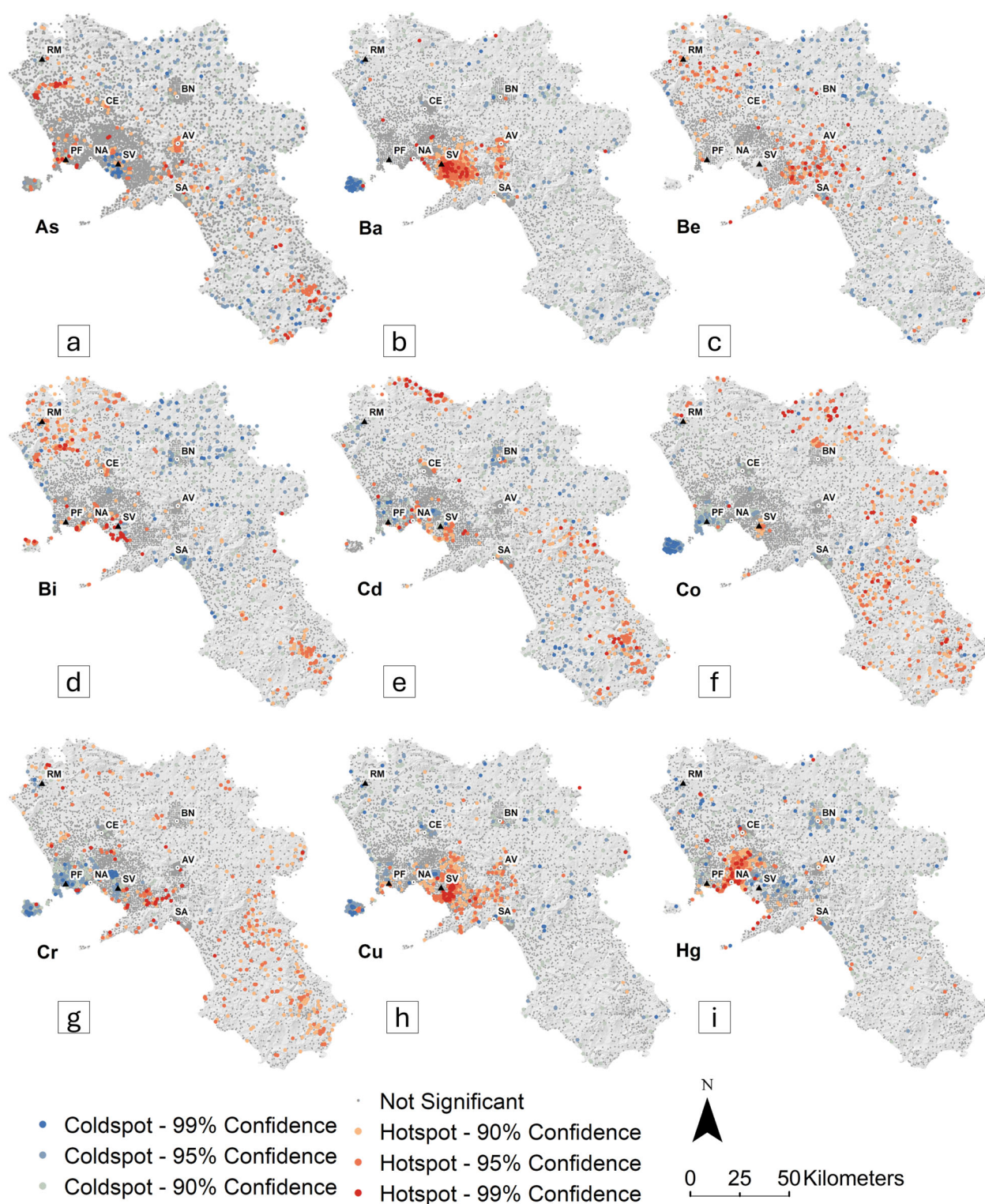


Figure 3. Spatial distribution of statistically significant hot- and coldspots for potentially toxic elements across the study area. Cold- and hotspots are shown at 90%, 95%, and 99% confidence levels using the Getis–Ord G_i^* statistic. Grey areas indicate locations without statistically significant clustering. Volcanic provinces: RM (Roccamonfina) and PF (Phlegraean Fields). City centers: NA (Naples), SV (Somma–Vesuvius), CE (Caserta), BN (Benevento), AV (Avellino), SA (Salerno). Scale bar: 25 km. Panels correspond to the following elements: (a) As, (b) Ba, (c) Be, (d) Bi, (e) Cd, (f) Co, (g) Cr, (h) Cu, and (i) Hg.

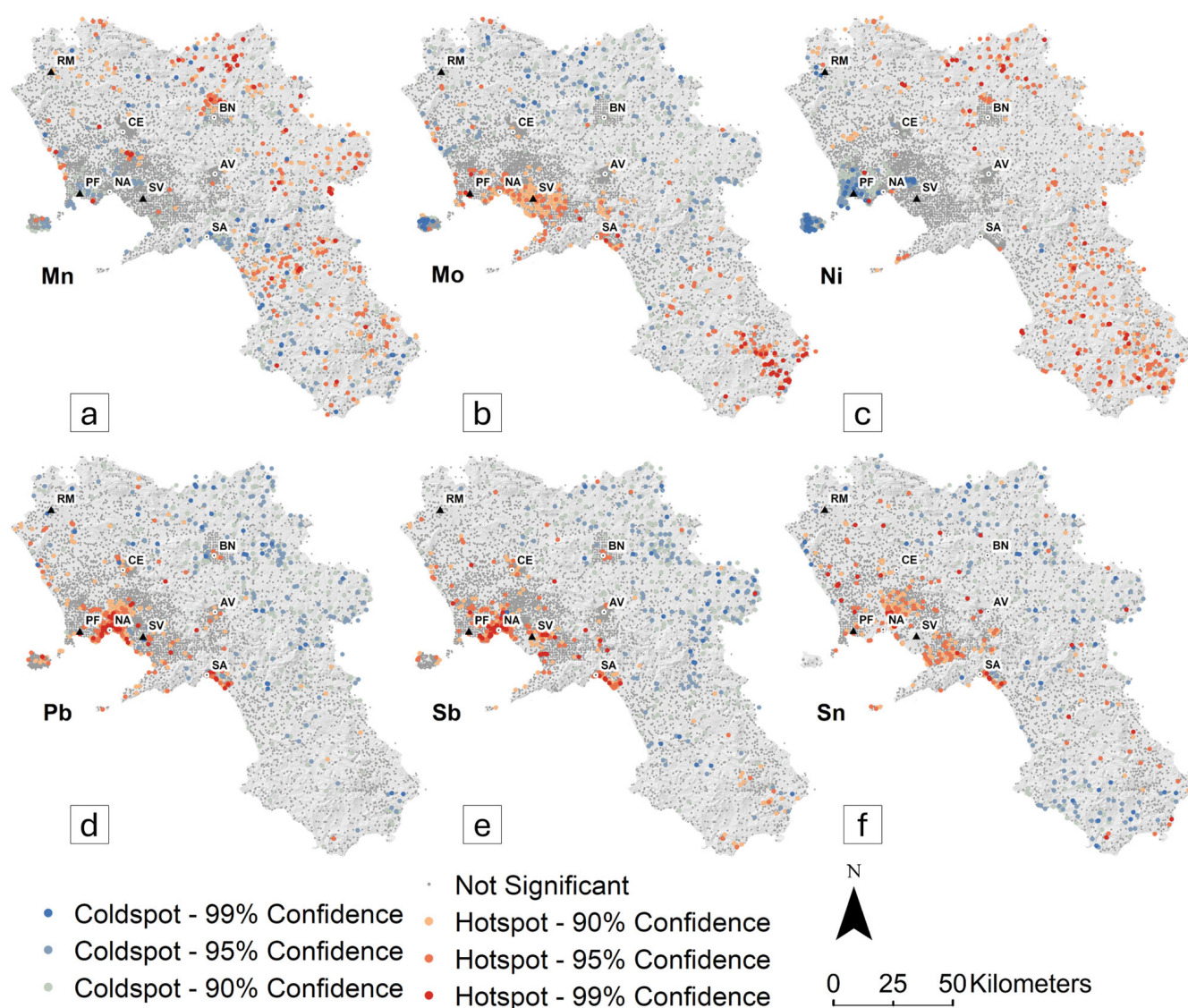


Figure 4. Spatial distribution of statistically significant hot- and coldspots for potentially toxic elements across the study area. Cold- and hotspots are shown at 90%, 95%, and 99% confidence levels using the Getis–Ord G_i^* statistic. Grey areas indicate locations without statistically significant clustering. Volcanic provinces: RM (Roccamonfina) and PF (Phlegraean Fields). City centers: NA (Naples), SV (Somma–Vesuvius), CE (Caserta), BN (Benevento), AV (Avellino), SA (Salerno). Scale bar: 25 km. Panels correspond to the following elements: (a) Mn, (b) Mo, (c) Ni, (d) Pb, (e) Sb, and (f) Sn.

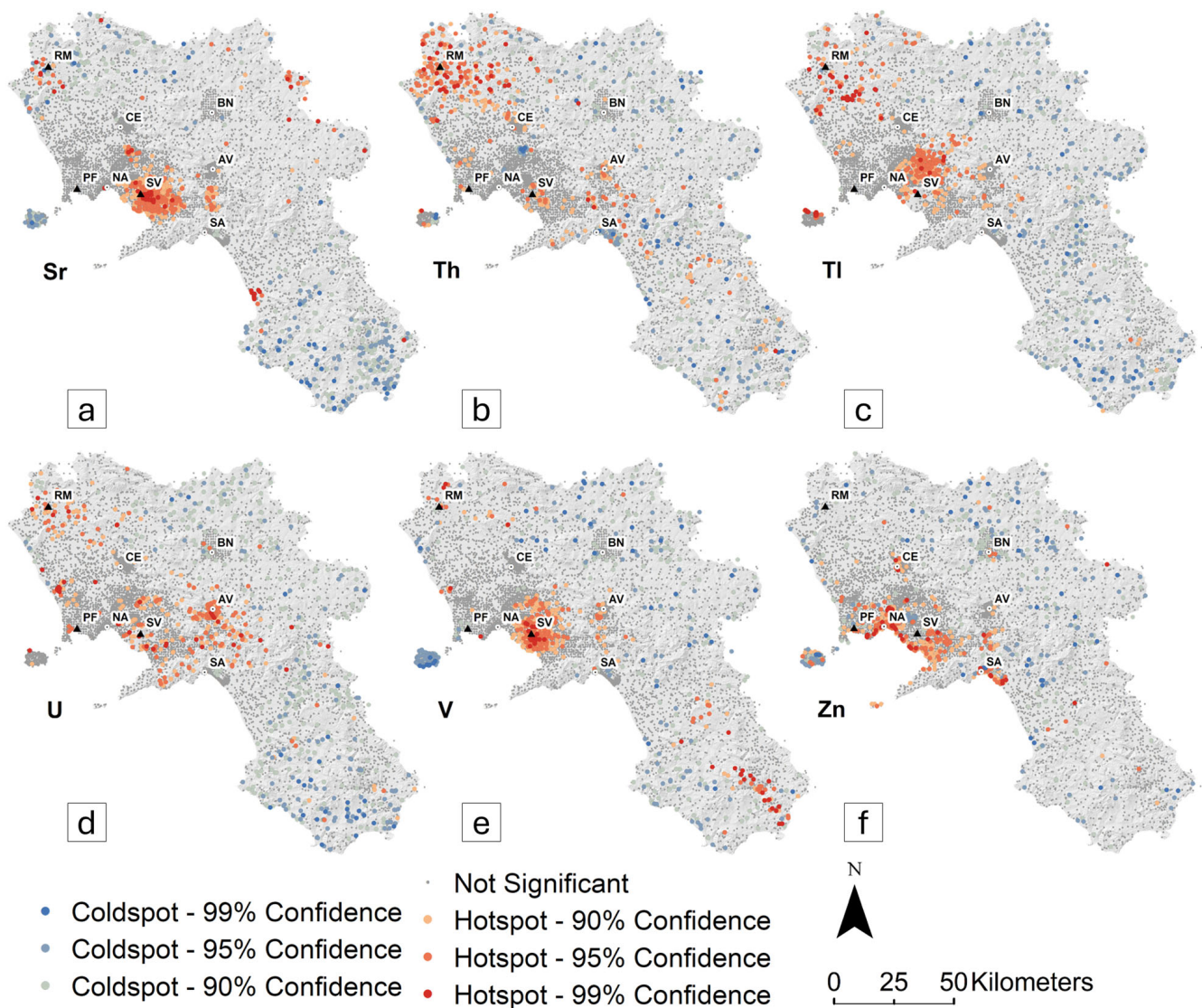


Figure 5. Spatial distribution of statistically significant hot- and coldspots for potentially toxic elements across the study area. Cold- and hotspots are shown at 90%, 95%, and 99% confidence levels using the Getis–Ord G_i^* statistic. Grey areas indicate locations without statistically significant clustering. Volcanic provinces: RM (Roccamonfina) and PF (Phlegraean Fields). City centers: NA (Naples), SV (Somma–Vesuvius), CE (Caserta), BN (Benevento), AV (Avellino), SA (Salerno). Scale bar: 25 km. Panels correspond to the following elements: (a) Sr, (b) Th, (c) Tl, (d) U, (e) V, and (f) Zn.

4. Discussion

4.1. Hotspot Clustering and Controlling Factors

The broad data dispersion, significant skewness, and presence of outliers in the raw dataset, as indicated by the statistical indices, reflect a non-normal distribution, which is commonly observed in cases of both geological heterogeneity and localized contamination [12]. The application of the NST effectively converted the data into a distribution approximating the Gaussian distribution by assigning values based on their rank within a normal distribution. The NST offers the advantage of a non-linear, rank-based adjustment that enhances the suitability of variance analyses [33]. This approach also reduces the impact of outliers by assigning them a lower statistical influence than central values. Supplementary Material (Figures S1 and S2) illustrates the effect of the NST. Histograms with fitted density curves of normal score-transformed concentrations showed that the

originally skewed distributions are converted to approximately standard normal form (mean = 0, variance = 1), thereby allowing robust geostatistical processing.

Hotspot analysis confirmed geology as the main controlling factor governing the elemental clustering patterns (see Figures 3–5). Supplementary Material (Table S1) provides an environmental and regulatory framework, reporting the Contamination Threshold Concentrations established by the Italian Legislative Decree 152/2006. These values represent the contamination limits for soils based on use (residential/green areas and industrial/commercial sites) and define the concentration levels above which a site is classified as potentially contaminated and requires further investigation. Threshold concentrations are not established in the current legislation for Ba, Bi, Mn, Mo, Sr, Th, and U.

Specifically, As, Be, Bi, Th, Tl, and U were associated with soils developed from the weathering of potassic and ultrapotassic volcanic products of the Somma–Vesuvius and Roccamonfina volcanic systems [29]. The hotspots observed across the carbonate massifs highlighted distinct geochemical influences from volcanic fallout. In particular, the pyroclastic covers over the Sorrento Peninsula and the Lattari Mountains are associated with the 79 A.D. Plinian eruption of Vesuvius, whose widespread tephra dispersal affected coastal areas to the south [34]. Conversely, the more internal massifs of the Monti di Avellino and Monti Picentini, within the central Apennine belt, were predominantly influenced by distal tephra fallout from the Mercato Pumice eruption (9 ka) [35] and the Avellino eruption (4 ka) [34], both of which produced high columns and widespread ash and tephra affecting the eastern and southeastern sectors of the Campania region. In this framework, Be and Tl are of particular interest for environmental management, as they frequently exceed the contamination thresholds established by Italian legislation, despite their natural origin.

In addition, arsenic showed hotspots in correspondence with the Phlegraean Fields area, suggesting a relation with the volcanic deposits and/or active hydrothermal activity [36,37]. The diffuse degassing and fluid–soil interactions typical of active hydrothermal systems may promote the mobilization and accumulation of As in surface soils. This interpretation is consistent with studies reporting high As concentrations in hydrothermal fluids [38] from Phlegraean Fields, which support the role of gas emissions and shallow hydrothermal circulation in controlling arsenic distribution in the environment. The spatial clustering of Ba, Cu, and Sr appeared concentrated across the Somma–Vesuvius area, highlighting the geochemical signature of the younger volcanic products derived from the latest Vesuvius eruptions. The enrichment of these elements may be associated with the gravitational separation of minerals, such as augite, during magma differentiation, as well as carbonate assimilation processes [39]. Similarly, the strong spatial clustering of V in the Somma–Vesuvius area likely reflects the geochemical fingerprint of volcanic products associated with the most recent eruptive phases. As highlighted by Melluso et al. [40], vanadium shows a strong affinity for garnet, a mineral phase occurring in these volcanic rocks, and clinopyroxene. In the case of Cu, however, distinguishing between geogenic and anthropogenic sources in topsoils is a significant challenge in environmental geochemistry, as its geochemical signal may also reflect anthropogenic contributions from the use of copper-based pesticides and fertilizers in agricultural practices [41–44]. Indeed, this area is extensively cultivated with vineyards, orchards, and vegetables, where Cu compounds are commonly applied for phytosanitary treatments [45,46]. The spatial clustering patterns of Co, Cr, Ni, and Mn showed strong correspondence with the region's siliciclastic units. This behavior is controlled by a common geochemical mechanism, as Cr, Co, and Ni are known to co-precipitate with Mn oxides. These manganese-bearing phases are more abundant in finer soil particles, which provide a higher specific surface area and enhanced capacity for metal adsorption and retention [47]. This process may explain the observed co-enrichment of these transition metals in soils derived from siliciclastic lithotypes. Regarding Cr, its

numerous industrial applications make it particularly widespread in the environment. The area between the Somma–Vesuvius volcano, Avellino, and Salerno borders the Sarno River basin, which is historically recognized for elevated concentrations of chromium of anthropogenic origin. This enrichment is mainly attributed to leather tanning, where chromium tanning is used to produce durable, soft, and color-resistant leather products [48,49]. The similar spatial clustering of Mo, Sn, Pb, Sb, Cd, Zn, and Hg is strongly associated with the urbanized and industrialized areas of the Campania region. These elements exhibit elevated concentrations in correspondence with major city centers, especially Napoli and Salerno, as well as areas of intense anthropogenic pressure. Their distribution highlighted the influence of long-standing human activities, including industrial processes, vehicular traffic, and intensive agriculture [50,51]. These findings are consistent with the behavior of these metals as typical tracers of anthropogenic input, often linked to emissions, waste disposal, and the use of agrochemicals [52,53]. However, for Cd, a spatial correspondence is observed between hotspots and the occurrence of carbonatic units in the study area. This pattern suggests that, in addition to anthropogenic inputs, Cd may also be influenced by natural factors, including weathering and pedogenetic processes that affect the carbonatic substrate, which locally contribute to Cd accumulation in topsoils [54].

4.2. Data Filtering and Regional Geochemical Baseline Estimation

The methodological approach adopted for baseline estimation of the six PTEs of anthropogenic interest (Zn, Sb, Hg, Pb, Sn, and Cd) yielded an efficient isolation of localized anomalies associated with site-specific enrichment. By excluding all statistically significant hotspots ($Gi^* = 1, 2, 3$), the proposed workflow removed data points potentially linked to localized contamination sources such as industrial activities, traffic emissions, and agricultural practices, as well as natural localized enrichments, exemplified by certain Cd hotspots. This selective removal enhanced the reliability of the regional baseline calculation, ensuring that the results were not biased by extreme or spatially clustered high values. Subdividing the dataset by lithologic type (alluvial, carbonatic, siliciclastic, volcanic) further strengthened the accuracy of baseline values by highlighting the role of the parent material in determining the natural concentrations of elements in soils. The application of the Box–Cox transformation before statistical treatment ensured the normalization of each lithological subset, thereby improving the reliability of UTL95% with 95% coverage (UTL95-95) estimation and satisfying the statistical assumptions required for this type of percentile-based baseline determination.

To provide a consistent comparison, baseline values derived by De Vivo et al. [28] and background values of Pacifico et al. [32] were also considered. In the first case, these values, extracted from the same original dataset, were obtained using the Spectrum–Area (S-A) analysis, a data-driven approach based on frequency that distinguishes between background/baseline values, the transitional zone (noise), and anomalies. The S-A method analyses the frequency spectrum of the geochemical signal and identifies characteristic thresholds by evaluating the relationship between the spectral energy density and the spatial distribution (area) of the signal [55,56]. Specifically, the method detects inflexion points where the slope of the frequency signal changes significantly, marking the transition from geochemically homogeneous background/baseline populations to anomalous values potentially related to external inputs. After isolating the portion of data classified as geochemical baseline, the UTL95-95 was calculated using the same statistical procedure applied to our Gi^* Bin-based baseline extraction, allowing the two approaches to be compared under similar conditions.

In Pacifico et al. [32], the background values were obtained using the geochemical gene concept, which employs relatively immobile elements (Fe, U, Th, La, Ti, Al) to define natural

soils conditions, and specific indicator elements to trace human impact (Pb, Zn, Au, Sb, Hg for urban sources; Cu, As, P, Na, K for agricultural ones). Each topsoil sample was compared with an “ideal gene” code, generating a similarity index that quantifies contamination from 0% (highly contaminated) to 100% (uncontaminated). Samples with a similarity index between 80% and 100% were considered representative of natural conditions or background and were therefore used to calculate the UTL95-90, ensuring that background values were derived exclusively from soils unaffected by anthropogenic inputs.

To facilitate an immediate comparison between the different approaches, baseline ratios based on the UTL values reported by De Vivo et al. [28] and Pacifico et al. [32] and those from this study were calculated and displayed in Figure 6 as a complementary graphical representation.

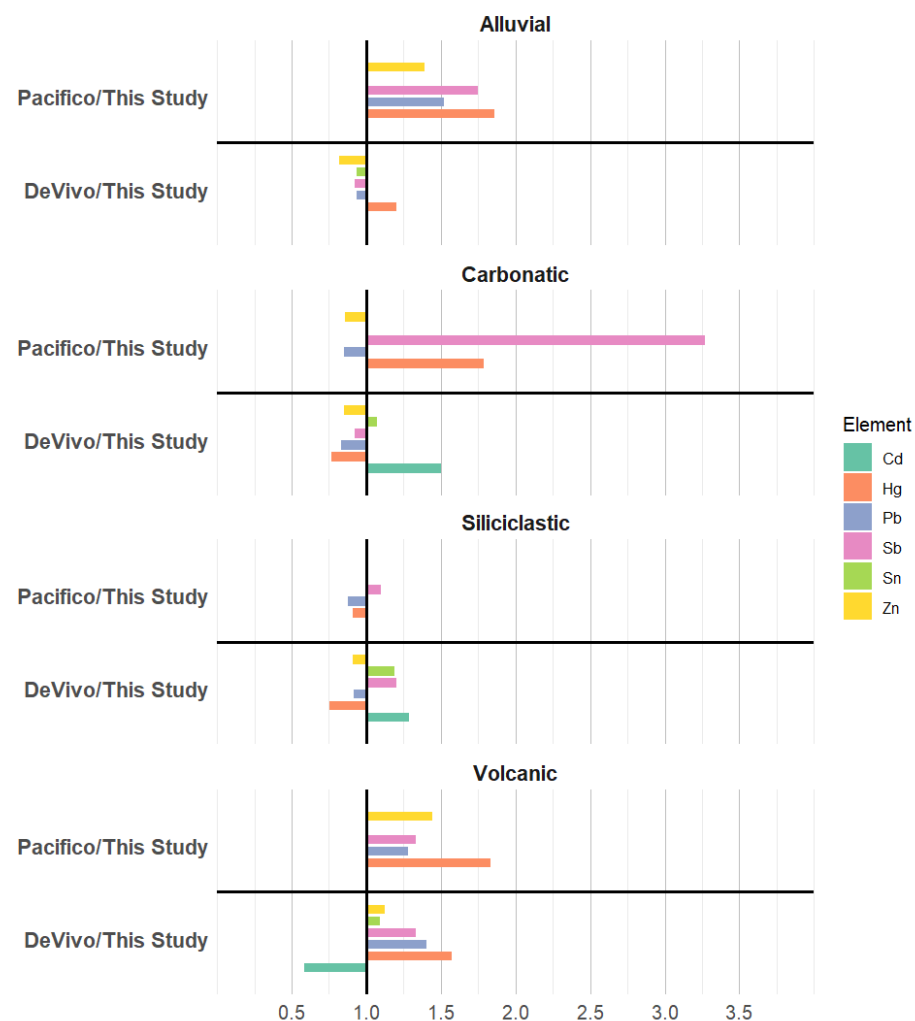


Figure 6. Baseline ratios between threshold values identified by De Vivo et al. [28] and Pacifico et al. [32] and those estimated in this study, calculated for each lithological unit (alluvial, carbonatic, siliciclastic, and volcanic).

Figure 6 highlights the magnitude of the differences between baseline values across the lithological units (alluvial, carbonatic, siliciclastic, and volcanic). Values close to unity indicate good agreement between methods, whereas ratios greater than 1 indicate higher baseline values reported by De Vivo et al. [28] or Pacifico et al. [32] than those reported in this study; conversely, ratios below 1 indicate cases in which the baseline values estimated in this study are higher than those reported in the literature.

The baseline values derived in this study show comparable trends in elemental variability across the different lithological units, consistent with both De Vivo et al. [28] and

Pacifico et al. [32]. Compared with De Vivo et al. [28], the baseline ratios are generally similar or slightly lower, with few exceptions where higher values were observed in De Vivo et al. [28], such as for Hg in the alluvial and volcanic units and Pb in the volcanic unit. In comparison, lower values characterize Hg in the siliciclastic and Zn in the alluvial units. In contrast, the comparison with Pacifico et al. [32] shows several differences, reporting generally higher background concentrations across most lithologies, consistent with the inclusion of a broader geochemical variability within their similarity-based model. This is particularly evident for Hg in volcanic soils, which show baseline values almost twice those estimated in this study. Since the volcanic domain largely coincides with the urbanized area of the city of Naples, this discrepancy likely reflects incomplete exclusion of anomalous concentrations from anthropogenic sources, thereby inflating the estimated threshold values. Conversely, Sb stands out for exhibiting values approximately 3.5 times higher within the carbonatic unit. However, this result should be interpreted with caution, as the carbonatic domain considered by Pacifico et al. [32] is substantially more limited than that adopted in the present study, as it included carbonatic units with pyroclastic coverage as part of the volcanic domain. The reduced spatial extent and sample size are likely to amplify the influence of locally elevated Sb concentrations, resulting in a statistically inflated baseline value. Nevertheless, some elements in Pacifico et al. are lower than in this study, particularly Pb in the carbonate and siliciclastic units, Hg in the siliciclastic unit, and Zn in the carbonate unit.

These discrepancies are likely due to conceptual and methodological differences underlying the methods. De Vivo et al. [28] and Pacifico et al. [32] used two data-driven approaches that rely on statistical differentiation of data distributions and on the quantification of background levels through the concept of similarity, respectively, independently of the dataset's spatial structure. In contrast, the hotspot-based approach adopted here integrates spatial context, excluding clustered anomalies that are potentially anthropogenic, such as industrial activity, traffic emissions, or agricultural contamination, and natural enrichments, thus providing less conservative and spatially consistent baseline estimates.

4.3. Limits, Implications, and Future Research

The proposed approach is methodologically straightforward, requiring a geochemically consistent topsoil dataset, data normalization (e.g., NST or Box-Cox transformation), and subsequent hotspot analysis. Its applicability can be extended to other regions, provided that sampling density and spatial resolution are appropriate for the scale of investigation. However, removing hotspots may introduce bias in background estimation, particularly for elements subject to natural geogenic enrichment (e.g., Be, As, Tl). In such cases, the procedure may lead to an underestimation of site-specific baseline levels by discarding naturally elevated concentrations. This appears to be the case for Cd, where the applied filtering may have attenuated both anthropogenic and natural signatures. Moreover, the outcome of the hotspot analysis is highly sensitive to parameters such as neighborhood definition, weight matrix specification, and search radius, which govern the detection of statistically significant clusters.

Although the method performed adequately at the regional scale considered in this study, it did not substantially enhance source discrimination. Future applications should therefore consider coupling hotspot spatial patterns with complementary datasets, such as land-use information, detailed geological mapping, and multivariate statistics, to improve the resolution and interpretability of source attribution analyses.

5. Conclusions

The spatial distribution of hotspots identified in this study shows strong consistency with the multivariate patterns previously described by Iannone et al. [18] while providing the additional advantage of allowing a detailed, element-specific interpretation rather than focusing solely on multivariate associations. The observed geochemical patterns confirm that As, Be, Bi, Th, Tl, U, and V are mainly associated with volcanic lithologies. In contrast, Co, Cr, Mn, and Ni reflect the influence of siliciclastic units, and Cd, Hg, Pb, Sb, Sn, and Zn are predominantly linked to anthropogenic activities. For the six PTEs used in baseline estimation, the exclusion of hotspots ($G_i^* \text{ Bin} \geq 1$) effectively reduced the influence of local enrichments, facilitating the definition of clearer regional trends. However, in the case of Cd, this procedure may also have inadvertently removed natural geogenic signals. Overall, the resulting baseline values are generally consistent with those from previous studies in the same area, with only minor local deviations attributable to site-specific factors.

Future developments should aim to integrate hotspot-based analysis with more detailed supporting data, including detailed geology, land-use data, and multivariate approaches. Such integration would enhance the capacity to discriminate between natural and anthropogenic sources, thereby improving baseline definition and minimizing the unintended removal of natural variability.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/environments13020098/s1>, Figure S1: Histograms with fitted density curves of normal score-transformed concentrations of selected trace elements (As, Ba, Be, Bi, Cd, Co, Cr, Cu, Hg, Mn, Mo, Ni); Figure S2: Histograms with fitted density curves of normal score-transformed concentrations of selected trace elements (Pb, Sb, Sn, Sr, Th, Tl, U, V, Zn). Table S1: Contamination threshold concentrations for the selected elements according to the Italian Legislative Decree 152/2006.

Author Contributions: Conceptualization, A.I. and S.D.; methodology, A.I., C.Z. and S.A.; software, A.I.; validation, S.D.; formal analysis, A.I.; investigation, A.I. and S.D.; resources, S.A.; data curation, A.I. and S.D.; writing—original draft preparation, A.I. and S.D.; writing—review and editing, S.A.; visualization, A.I.; supervision, S.A. and C.Z.; project administration, S.A.; funding acquisition, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This study was carried out within the RETURN Extended Partnership and received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan—NRRP, Mission 4, Component 2, Investment 1.3—D.D. 1243 2/8/2022, PE0000005).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

K-S	Kolmogorov–Smirnov
NST	Normal Score Transformation
PTE	Potential Toxic Element
UTL	Upper Tolerane Limit

References

1. Tipping, E.; Lawlor, A.J.; Loftis, S.; Shotbolt, L. Simulating the Long-Term Chemistry of an Upland UK Catchment: Heavy Metals. *Environ. Pollut.* **2006**, *141*, 139–150. [[CrossRef](#)]
2. Hossain, M.M.; Jahan, I.; Dar, M.A.; Dhanavade, M.J.; Mamtaz, A.F.B.; Maxwell, S.J.; Han, S.; Zhu, D. A Review of Potentially Toxic Elements in Sediment, Water, and Aquatic Species from the River Ecosystems. *Toxics* **2024**, *13*, 26. [[CrossRef](#)]

3. Alloway, B. *Heavy Metals in Soils: Trace Metals and Metalloids in Soils and Their Bioavailability*; Springer: Dordrecht, The Netherlands, 2013.
4. Zhang, J.; Peng, W.; Lin, M.; Liu, C.; Chen, S.; Wang, X.; Gui, H. Environmental Geochemical Baseline Determination and Pollution Assessment of Heavy Metals in Farmland Soil of Typical Coal-Based Cities: A Case Study of Suzhou City in Anhui Province, China. *Heliyon* **2023**, *9*, e14841. [[CrossRef](#)]
5. Borůvka, L.; Vacek, O.; Jehlička, J. Principal Component Analysis as a Tool to Indicate the Origin of Potentially Toxic Elements in Soils. *Geoderma* **2005**, *128*, 289–300. [[CrossRef](#)]
6. Zhang, H.; Yu, M.; Xu, H.; Wen, H.; Fan, H.; Wang, T.; Liu, J. Geochemical Baseline Determination and Contamination of Heavy Metals in the Urban Topsoil of Fuxin City, China. *J. Arid Land* **2020**, *12*, 1001–1017. [[CrossRef](#)]
7. Liu, J.; Gong, C.; Tan, C.; Wen, L.; Li, Z.; Liu, X.; Yang, Z. Geochemical Baseline Establishment and Accumulation Characteristics of Soil Heavy Metals in Sabaochaqu Watershed at the Source of Yangtze River, Qinghai-Tibet Plateau. *Sci. Rep.* **2024**, *14*, 21945. [[CrossRef](#)]
8. Salminen, R.; Gregorauskien, V. Considerations Regarding the Definition of a Geochemical Baseline of Elements in the Surficial Materials in Areas Differing in Basic Geology. *Appl. Geochem.* **2000**, *15*, 647–653. [[CrossRef](#)]
9. Matschullat, J.; Ottenstein, R.; Reimann, C. Geochemical Background—Can We Calculate It? *Environ. Geol.* **2000**, *39*, 990–1000. [[CrossRef](#)]
10. Reimann, C.; Garrett, R.G. Geochemical Background—Concept and Reality. *Sci. Total Environ.* **2005**, *350*, 12–27. [[CrossRef](#)]
11. USEPA. *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*; Office of Emergency and Remedial Response U.S. Environmental Protection Agency: Washington, DC, USA, 2002.
12. Reimann, C.; Filzmoser, P.; Garrett, R.; Dutter, R. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
13. Tukey, J.W. Exploratory Data Analysis. In *The Concise Encyclopedia of Statistics*; Springer: New York, NY, USA, 2008; pp. 192–194.
14. Cheng, Q.; Agterberg, F.P.; Ballantyne, S.B. The Separation of Geochemical Anomalies from Background by Fractal Methods. *J. Geochem. Explor.* **1994**, *51*, 109–130. [[CrossRef](#)]
15. Cheng, Q.; Xu, Y.; Grunsky, E. Integrated Spatial and Spectrum Method for Geochemical Anomaly Separation. *Nat. Resour. Res.* **2000**, *9*, 43–52. [[CrossRef](#)]
16. Zuo, R.; Wang, J. Fractal/Multifractal Modeling of Geochemical Data: A Review. *J. Geochem. Explor.* **2016**, *164*, 33–41. [[CrossRef](#)]
17. Nguyen, T.T.; Vu, T.D. Identification of Multivariate Geochemical Anomalies Using Spatial Autocorrelation Analysis and Robust Statistics. *Ore Geol. Rev.* **2019**, *111*, 102985. [[CrossRef](#)]
18. Iannone, A.; Dominech, S.; Zhang, C.; Pacifico, L.R.; De Falco, A.; Albanese, S. Principal Component Analysis to Discriminate and Locate Natural and Anthropogenic Sources of Contamination Within a Strongly Anthropized Region: A Technical Workflow. *Environments* **2025**, *12*, 163. [[CrossRef](#)]
19. Ord, J.K.; Getis, A. Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geogr. Anal.* **1995**, *27*, 286–306. [[CrossRef](#)]
20. Xu, H.; Demetriades, A.; Reimann, C.; Jiménez, J.J.; Filser, J.; Zhang, C. Identification of the Co-Existence of Low Total Organic Carbon Contents and Low PH Values in Agricultural Soil in North-Central Europe Using Hot Spot Analysis Based on GEMAS Project Data. *Sci. Total Environ.* **2019**, *678*, 94–104. [[CrossRef](#)] [[PubMed](#)]
21. Giannetti, B.; Luhr, J.F. The White Trachytic Tuff of Roccamonfina Volcano (Roman Region, Italy). *Contrib. Mineral. Petrol.* **1983**, *84*, 235–252. [[CrossRef](#)]
22. Peccerillo, A. *Plio-Quaternary Volcanism in Italy: Petrology, Geochemistry, Geodynamics*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1–365. [[CrossRef](#)]
23. De Vivo, B.; Rolandi, G.; Gans, P.; Calvert, A.; Bohrsen, W.; Spera, F.; Belkin, H. New Constraints on the Pyroclastic Eruptive History of the Campanian Volcanic Plain (Italy). *Mineral. Petrol.* **2001**, *73*, 47–65. [[CrossRef](#)]
24. Rolandi, G.; Petrosino, P.; Mc Geehin, J. The Interplinian Activity at Somma-Vesuvius in the Last 3500 Years. *J. Volcanol. Geotherm. Res.* **1998**, *82*, 19–52. [[CrossRef](#)]
25. Rolandi, G.; Bellucci, F.; Heizler, M.T.; Belkin, H.E.; De Vivo, B. Tectonic Controls on the Genesis of Ignimbrites from the Campanian Volcanic Zone, Southern Italy. *Mineral. Petrol.* **2003**, *79*, 3–31. [[CrossRef](#)]
26. Cicchella, D.; Ambrosino, M.; Albanese, S.; Guarino, A.; Lima, A.; De Vivo, B.; Guagliardi, I. Major Elements Concentration in Soils. A Case Study from Campania Region (Italy). *J. Geochem. Explor.* **2023**, *247*, 107179. [[CrossRef](#)]
27. Salminen, R.; Batista, M.; Bidovec, M.; Demetriades, A.; De Vivo, B.; De Vos, W.; Duris, M.; Gilucis, A.; Gregorauskien, V.; Halamić, J.; et al. *FOREGS Geochemical Atlas of Europe, Part 1: Background Information, Methodology and Maps*; Geological Survey of Finland GTK: Espoo, Finland, 2005.
28. De Vivo, B.; Cicchella, D.; Lima, A.; Fortelli, A.; Guarino, A.; Zuzolo, D.; Esposito, M.; Cerino, R.; Pizzolante, A.; Albanese, S. *Monitoraggio Geochimico-Ambientale dei Suoli della Regione Campania—Progetto Campania Trasparente—Volume 1—Elementi Potenzialmente Tossici e Loro Biodisponibilità, Elementi Maggiori e in Traccia (Distribuzione in Suoli Superficiali e Profondi)*; Aracne Editrice: Rome, Italy, 2021.

29. Zuzolo, D.; Cicchella, D.; Lima, A.; Guagliardi, I.; Cerino, P.; Pizzolante, A.; Thiombane, M.; De Vivo, B.; Albanese, S. Potentially Toxic Elements in Soils of Campania Region (Southern Italy): Combining Raw and Compositional Data. *J. Geochem. Explor.* **2020**, *213*, 106524. [\[CrossRef\]](#)
30. Getis, A.; Ord, J.K. The Analysis of Spatial Association by Use of Distance Statistics. *Geogr. Anal.* **1992**, *24*, 189–206. [\[CrossRef\]](#)
31. Box, G.E.P.; Cox, D.R. An Analysis of Transformations. *J. R. Stat. Soc. Ser. B (Methodol.)* **1964**, *26*, 211–243. [\[CrossRef\]](#)
32. Pacifico, L.R.; Carotenuto, F.; Guarino, A.; Iannone, A.; Cicchella, D.; Albanese, S. The Anthropigene: A New Approach in Environmental Geochemistry to Discriminate Anomalies from Natural Background. *J. Geochem. Explor.* **2025**, *278*, 107832. [\[CrossRef\]](#)
33. Ludwig, O. Blom, Gunnar: Statistical Estimates and Transformed Beta-variables. Wiley/New York, Almquist Und Wiksell/Stockholm 1958; 176 S., Kr. 20,—. *Biom. Z.* **1961**, *3*, 285. [\[CrossRef\]](#)
34. Ayuso, R.A.; De Vivo, B.; Rolandi, G.; Seal II, R.R.; Paone, A. Geochemical and Isotopic (Nd-Pb-Sr-O) Variations Bearing on the Genesis of Volcanic Rocks from Vesuvius, Italy. *J. Volcanol. Geotherm. Res.* **1998**, *82*, 53–78. [\[CrossRef\]](#)
35. Mele, D.; Sulpizio, R.; Dellino, P.; La Volpe, L. Stratigraphy and eruptive dynamics of a pulsating Plinian eruption of Somma-Vesuvius: The Pomici di Mercato (8900 years B.P.). *Bull. Volcanol.* **2011**, *73*, 257–278. [\[CrossRef\]](#)
36. Masuda, H. Arsenic Cycling in the Earth's Crust and Hydrosphere: Interaction between Naturally Occurring Arsenic and Human Activities. *Prog. Earth Planet. Sci.* **2018**, *5*, 68. [\[CrossRef\]](#)
37. Tabelin, C.B.; Igarashi, T. Mechanisms of Arsenic and Lead Release from Hydrothermally Altered Rock. *J. Hazard. Mater.* **2009**, *169*, 980–990. [\[CrossRef\]](#)
38. Aiuppa, A.; D'Alessandro, W.; Federico, C.; Palumbo, B.; Valenza, M. The Aquatic Geochemistry of Arsenic in Volcanic Groundwaters from Southern Italy. *Appl. Geochem.* **2003**, *18*, 1283–1296. [\[CrossRef\]](#)
39. Savelli, C. The Problem of Rock Assimilation by Somma-Vesuvius Magma. *Contrib. Mineral. Petrol.* **1967**, *16*, 328–353. [\[CrossRef\]](#)
40. Melluso, L.; Scarpato, C.; Zanetti, A.; Sparice, D.; de' Gennaro, R. The Petrogenesis of Chemically Zoned, Phonolitic, Plinian and Sub-Plinian Eruptions of Somma-Vesuvius, Italy: Role of Accessory Phase Removal, Independently Filled Magma Reservoirs with Time, and Transition from Slightly to Highly Silica Undersaturated Magmatic Series in an Ultrapotassic Stratovolcano. *Lithos* **2022**, *430–431*, 106854. [\[CrossRef\]](#)
41. Gimeno-García, E.; Andreu, V.; Boluda, R. Heavy Metals Incidence in the Application of Inorganic Fertilizers and Pesticides to Rice Farming Soils. *Environ. Pollut.* **1996**, *92*, 19–25. [\[CrossRef\]](#)
42. Karimi, B.; Masson, V.; Guiland, C.; Leroy, E.; Pellegrinelli, S.; Giboulot, E.; Maron, P.-A.; Ranjard, L. Ecotoxicity of Copper Input and Accumulation for Soil Biodiversity in Vineyards. *Environ. Chem. Lett.* **2021**, *19*, 2013–2030. [\[CrossRef\]](#)
43. Reimann, C.; Fabian, K.; Flem, B.; Englmaier, P. The Large-Scale Distribution of Cu and Zn in Sub- and Topsoil: Separating Topsoil Bioaccumulation and Natural Matrix Effects from Diffuse and Regional Contamination. *Sci. Total Environ.* **2019**, *655*, 730–740. [\[CrossRef\]](#)
44. Křibek, B.; Majer, V.; Veselovský, F.; Nyambe, I. Discrimination of Lithogenic and Anthropogenic Sources of Metals and Sulphur in Soils of the Central-Northern Part of the Zambian Copperbelt Mining District: A Topsoil vs. Subsurface Soil Concept. *J. Geochem. Explor.* **2010**, *104*, 69–86. [\[CrossRef\]](#)
45. Minolfi, G.; Petrik, A.; Albanese, S.; Lima, A.; Cannatelli, C.; Rezza, C.; De Vivo, B. The Distribution of Pb, Cu and Zn in Topsoil of the Campanian Region, Italy. *Geochem. Explor. Environ. Anal.* **2019**, *19*, 205–215. [\[CrossRef\]](#)
46. Cicchella, D.; De Vivo, B.; Lima, A. Background and Baseline Concentration Values of Elements Harmful to Human Health in the Volcanic Soils of the Metropolitan and Provincial Areas of Napoli (Italy). *Geochem. Explor. Environ. Anal.* **2005**, *5*, 29–40. [\[CrossRef\]](#)
47. Mistikawy, J.A.; Mackowiak, T.J.; Butler, M.J.; Mischenko, I.C.; Cernak, R.S.; Richardson, J.B. Chromium, Manganese, Nickel, and Cobalt Mobility and Bioavailability from Mafic-to-Ultramafic Mine Spoil Weathering in Western Massachusetts, USA. *Environ. Geochem. Health* **2020**, *42*, 3263–3279. [\[CrossRef\]](#)
48. Arienzo, M.; Adamo, P.; Rosaria Bianco, M.; Violante, P. Impact of Land Use and Urban Runoff on the Contamination of the Sarno River Basin in Southwestern Italy. *Water Air Soil Pollut.* **2001**, *131*, 349–366. [\[CrossRef\]](#)
49. Bini, C.; Maleci, L.; Romanin, A. The Chromium Issue in Soils of the Leather Tannery District in Italy. *J. Geochem. Explor.* **2008**, *96*, 194–202. [\[CrossRef\]](#)
50. Kurt, M.A. Comparison of Trace Element and Heavy Metal Concentrations of Top and Bottom Soils in A Complex Land Use Area. *Carpathian J. Earth Environ. Sci.* **2018**, *13*, 47–56. [\[CrossRef\]](#)
51. Wang, X.-S.; Qin, Y. Some Characteristics of the Distribution of Heavy Metals in Urban Topsoil of Xuzhou, China. *Environ. Geochem. Health* **2007**, *29*, 11–19. [\[CrossRef\]](#)
52. Yu, H.; Ni, S.-J.; He, Z.-W.; Zhang, C.-J.; Nan, X.; Kong, B.; Weng, Z.-Y. Analysis of the Spatial Relationship between Heavy Metals in Soil and Human Activities Based on Landscape Geochemical Interpretation. *J. Geochem. Explor.* **2014**, *146*, 136–148. [\[CrossRef\]](#)
53. Senesil, G.S.; Baldassarre, G.; Senesi, N.; Radina, B. Trace Element Inputs into Soils by Anthropogenic Activities and Implications for Human Health. *Chemosphere* **1999**, *39*, 343–377. [\[CrossRef\]](#)

54. Xia, X.; Ji, J.; Yang, Z.; Han, H.; Huang, C.; Li, Y.; Zhang, W. Cadmium Risk in the Soil-Plant System Caused by Weathering of Carbonate Bedrock. *Chemosphere* **2020**, *254*, 126799. [[CrossRef](#)]
55. Cheng, Q. Spatial and Scaling Modelling for Geochemical Anomaly Separation. *J. Geochem. Explor.* **1999**, *65*, 175–194. [[CrossRef](#)]
56. Zuo, R.; Xia, Q.; Zhang, D. A Comparison Study of the C–A and S–A Models with Singularity Analysis to Identify Geochemical Anomalies in Covered Areas. *Appl. Geochem* **2013**, *33*, 165–172. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.