Environmental Earth Sciences (2025) 84:446
https://doi.org/10.1007/s12665-025-12458-z

ORIGINAL ARTICLE

Identification of coal combustion impacts on soil contamination by risk elements needs empirical holistic approach: case study in the Most Basin, Czech Republic

T. Matys Grygar[1,2] · S. Adamec[1] · Š. Tůmová[1,2] · H. Burdová[1] · S. Křížanecká[1] · D. Bretschneiderová[1] · M. Lučić[3] · I. Pavlů[4]

Received: 26 February 2025 / Accepted: 20 July 2025 / Published online: 29 July 2025

Abstract

This work promotes an empirical holistic approach to the reliable identification of anthropogenic contributions to elevated concentrations of risk elements in soils. It is demonstrated through the evaluation of the impact of coal-fired power plants on soils in the Most Basin in the Czech Republic, Central Europe. The origin of the coal seam in the Most Basin is inherently associated with the presence of geochemically anomalous rocks, including those with ore veins at the basin edges, which complicates the identification of human impacts. This study is based on analyses of risk elements (As, Be, Cd, Cu, Pb, Sb, Zn) and lithogenic elements (Al, Ca, Fe, K, Mn, Rb, Si, Sr, Ti, Zr) in approximately 1 m thick soil profiles, Bayes space methodology for analysis of their granulometric curves, analyses of polycyclic aromatic hydrocarbons (PAH) in topsoils, and empirical (verifiable and explainable) data mining. Risk element concentrations were subjected to principal component analysis (PCA) and multilinear regression with Al, Fe, Mn, K, Rb, Si, Ti, and/or Zr in aim to correct the results for natural variability of soils. The results demonstrated that basin floor is covered by a mosaic of sediments with varying lithogenic origin that are not specified (or are incorrectly specified) in geological maps. The nonlinearity of interelement relationship and considerable site-specificity of soil composition precluded quantification of risk element concentrations, but empirical data mining made it possible to evaluate a power plant impact on soils. Geogenic anomalies were found to be a dominant factor in the elevated concentrations of As and Pb (from felsic effusive rocks and mineralization), Cu (from mafic rocks), Sb (from mineralization), and Be (from felsic effusive rocks) in the basin soils, which have incorrectly been attributed to coal combustion in the recent past. PAH concentrations are the most straightforward indicators of the impact of coal combustion, along with soil contamination from Cd and Zn. The contamination related to coal combustion does not represent real toxicological risks in agricultural soils. The methodology employed in this work could be used to revisit previous studies that underestimated the natural complexity of soil chemistry in coal basins and paradigmatically exaggerated the impacts of coal combustion on soil risk elements.

Keywords Soil · Risk elements · Geogenic anomalies · Coal combustion

✉ S. Adamec
slavomir.adamec@ujep.cz

1 Faculty of Environment, J. E. Purkyně University in Ústí nad Labem, Ústí nad Labem, Czech Republic

2 Institute of Inorganic Chemistry, Czech Academy of Sciences, Řež 250 68, Czech Republic

3 Ruđer Bošković Institute, Division of Marine & Environmental Research, Bijenička Cesta 54, Zagreb 10000, Croatia

4 Department of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Olomouc 771 46, Czech Republic

Introduction

Research should deliver a realistic impact assessment of individual anthropogenic activities on the environment as a prerequisite for qualified societal decisions on future industrial strategies and landscape management. Exaggeration of human impacts on the environment could favour tendency to suppress any mining and industrial activities, that can endanger sustainable economic development. This concern also extends to energy production from coal and its possible consequences on soil contamination. Conventional geochemical soil mapping is not sufficient for distinguishing industrial impacts from geogenic anomalies (Adamec et al. 2024) for reasons specified in the subsequent paragraph. Cicchella et al. (2015), Ballabio et al. (2018, 2021), and Reimann et al. (2018) evaluated extensive European soil mapping datasets and found that geogenic anomalies and other natural factors (such as soil properties and vegetation) are responsible for a major part of the variability of soil risk elements, which hinders the identification of industrial impacts. Past anthropogenic activities, which can be suspected to have caused soil contamination, have frequently occurred in geogenically anomalous areas. For example, ore regions (mineralized areas) attracted metal smelting, and coal seams in sedimentary basins attracted energy production; both activities stimulated the development of further industries, increased population density, and thus enhanced human pressure on the environment. Historical mining and metallurgic activities have been dispersed over landscape in un-documented patterns (Hošek et al. 2024). The coincidence of human activities and natural anomalies, commonly stemming from local geology, makes distinguishing natural and anthropogenic contributions a challenging task (Ander et al. 2013; Armiento et al. 2022; Hošek et al. 2024), which requires a holistic approach (Baize and Sterckeman 2001; Sterckeman et al. 2006; Ballabio et al. 2018, 2021; Pampura et al. 2019; Adamec et al. 2024). However, a considerable portion of soil contamination studies still suffers from oversimplifications. A common paradigmatic association exists between coal mining and combustion and elevated contents of soil risk elements without explicit proof of a causal chain (Bhuiyan et al. 2010; George et al. 2015; Pandey et al. 2016; Hanousková et al. 2021; Skála et al. 2022; Rouhani et al. 2023; Jia et al. 2023; Feng et al. 2024). The demands of a holistic approach to soil geochemistry data are now bypassed by formal statistical analyses of geochemical datasets (Jia et al. 2023; Feng et al. 2024), neglecting the real factors controlling risk element concentrations in soils.

The prerequisites for unbiased soil contamination studies are known. Geological maps must be employed during the planning of fieldwork to reveal and address all geochemical specificities of the study area, such as the presence of certain volcanic rocks, bio- or chemogenic sediments, and local mineralization (Amorosi et al. 2014; Vilà and Martínez-Lladó 2015; Stafilov et al. 2018; Armiento et al. 2022; Adamec et al. 2024; García et al. 2024). Examination of relationships between soil risk and lithogenic elements is considerably helpful if the target area has varying bedrock geology (Cicchella et al. 2015; Matys Grygar et al. 2016), the soil lithogenic origin is spatially or vertically variable (Baize and Sterckeman 2001; Sterckeman et al. 2006; Waroszewski et al. 2017; Pampura et al. 2019; Aytop et al. 2023; Adamec et al. 2024; García et al. 2024; Hošek et al. 2024; Sun et al. 2024a, b), and soils contain variable percentages of sand and organic matter (McKinley et al. 2016; Ballabio et al. 2018). Soil depth profiles (Baize and Sterckeman 2001; Aytop et al. 2023; Adamec et al. 2024) or at least topsoil/subsoil pairs (Amorosi et al. 2014; Stafilov et al. 2018; García et al. 2024; Hošek et al. 2024) are indispensable because the local background concentrations of risk elements from deeper soil horizons can differ considerably from regional or global background and legislative limits (Amorosi et al. 2014; Aytop et al. 2023; García et al. 2024; Hošek et al. 2024). However, the majority of the preceding works on the impacts of coal combustion on soils have overlooked these principles.

Conventional geochemical topsoil mapping in the majority of national or continental soil monitoring surveys has been tailored to assess food production safety; thus, they have not been designed to distinguish between anthropogenic contamination and geogenic anomalies (Matys Grygar et al. 2023). This is evidenced by the equivocality in discussions of large soil monitoring datasets (Vácha et al. 2015; Ballabio et al. 2018; Reimann et al. 2018; Zhang et al. 2020a, b), which typically conclude that elevated concentrations of risk elements result from an unspecified combination of geogenic and anthropogenic causes. Neither the most sophisticated mathematical models, which use all available information sources, can avoid this equivocality (Zhang et al. 2020a, b; Skála et al. 2024). We assert that only targeted sampling and an understanding of local specificities can address this issue (Adamec et al. 2024).

The aim of this work is to separate the impact of coal combustion on soil risk elements in the Most Basin, Czech Republic, from natural soil compositional variability, serving as a real case that exemplifies the complexity of this task. The methodology used in this research can be recommended for any case studies of human impacts in geologically complicated areas that have recently been the focus of research (Tapia et al. 2012; Cicchella et al. 2015; Pandey et al. 2016; Murray et al. 2023; García et al. 2024; Sun et al. 2024a, b). The content of polycyclic aromatic hydrocarbons (PAH) was included as an independent measure of coal combustion's impact on the environment (Okedeyi et al. 2013; Li et al. 2020; Zhang J et al. 2020; Armiento et al. 2022; Kravchenko et al. 2024). The anthropogenic portion of the soil risk elements above their natural levels was calculated from lithogenic element concentrations, similarly to preceding studies (Sterckeman et al. 2006; Tapia et al. 2012; García et al. 2024). This work is distinguished by its exclusive use of white box (explainable, process-based, empirical) models derived from known geochemical principles, which are thus applicable to understanding the observed features in real-world terms rather than formal statistical measures. We prefer white box models to the increasingly popular black box models, which are suitable for mimicking datasets through algorithmic models with high mathematical precision but lack empirically verifiable physical or chemical bases, hypotheses, and conclusions. Our goal was to produce results that could meet forensic standards, as required by Miller (2013) for fluvial contamination studies.

This work corroborates previous study by Adamec et al. (2024) focused mostly on soils on sedimentary rocks in the coal basin floor. That preceding study has shown need to characterise in more details also soils on geogenically anomalous rocks from which a part of the basin floor sediments (and soils on them) have been formed. One of original aims of Adamec et al. (2024) was to quantify anthropogenic contamination by using some geochemical indices, but the conclusion was it is not successful. Novel reference area and extra sites were thus sampled for this work to make one more attempt to achieve this goal. This novel work additionally tests the preceding somehow surprising evaluation (Adamec et al. 2024) of rather weak impact of the coal combustion on soils in the basin floor. Newly added with respect to Adamec et al. (2024) were PAH concentrations as a proxy for the impacts of combustion of carbonaceous materials on soils.

Study area

Study area is the north-eastern part of the 80 km long Most Basin ended by city of Ústí nad Labem. Soft coal deposits were formed in the middle Miocene during early stages of basin floor deposition after the Oligocene volcanism in the Bohemian Massif triggered basin floor subsidence on response to the Alpine orogeny (Rajchl et al. 2009). Coal seam, typically 30 m thick, was formed from peat accumulated in a waste swamp on the basin floor (Rajchl et al. 2009). The earliest preserved note of coal mining dates back to 1403 in the town register of Duchcov (Otčenášek et al. 2019). Later the mining stepwise grew in particular during the Industrial Revolution; the largest expansion was after 1945, when surface mining of brown coal began to increase dramatically. Coal mining peaked in by the mid-1980s, the annual volume of coal extracted reached nearly one hundred million tons (Otčenášek et al. 2019).

The Most Basin is home to the major Czech power plants. Several coal-fired power plants and an energy complex are present in the part of the basin addressed in this work (Fig. 1). The primary focus here is on the potential impact of the Ledvice Power Plant, built in the 1960s, currently with an installed power of 770 MW. The reference area for this study is situated in the agricultural region southeast of the Most Basin (these results were taken from Adamec et al. 2024) and is separated from all basin power plants by the elevations of the České Středohoří Mountains (Fig. 1, the position of the reference area is indicated by green triangles).

The bedrock geology of the basin floor is spatially variable (Fig. S1 in the Supplementary Information). It includes Cretaceous marine sediments (marls), Miocene fluvial-lacustrine sediments, effusive felsic volcanic rocks (rhyolites, ignimbrites, and their tuffs), and Pleistocene unconsolidated sediments (loess, alluvial, and deluvial deposits). Major geogenic anomalies border the basin floor (Fig. 1). Paleozoic felsic intrusive (granite) and effusive (rhyolite, ignimbrite) and metamorphic rocks (mainly orthogneiss and paragneiss) outcrop mainly in the Ore Mountains along the northwest edge of the basin. Cenozoic mafic effusive volcanic rocks (basalt and basaltic tuffs) are found along the southeast edge of the basin and at occasional elevations within the basin floor. The risk element content in rocks is site specific (Matys Grygar et al. 2023) and thus “local geologically specific background values” cannot be defined for the study area (Adamec et al. 2024).

Metal ores were historically mined in the Ore Mountains; four ore regions occur not far from the basin floor in the study area (Fig. 1). The Czech Geological Survey (www.geology.cz) records abandoned mines and spoil heaps following the extraction of tin-tungsten ore, fluorite-barite veins, polymetallic ores, radioactive raw materials, pyrite, and other iron ores. The position of metallurgy centres and their particular emission impact has not been documented (Hošek et al. 2024). The soils in the Ore Mountains contain anomalously high concentrations of arsenic and lead, frequently exceeding legislative limits for agricultural soils (Vácha et al. 2015; Matys Grygar et al. 2023; Hošek et al. 2024). Those elevated concentrations result from geogenic and anthropogenic causes that can be distinguished at best as averages for larger areas (Hošek et al. 2024). The Ore Mountain soils are imported to the north-western part of the Most Basin floor also by local streams, as it is shown in the river network of the study area (Fig. S2). The downslope movement of these anomalous soils has also impacted the soils on the Most Basin floor and must be considered when evaluating the impact of coal combustion in the study area (Adamec et al. 2024).

Sampling and analyses

Soil sampling was conducted between 2021 and 2024 across ten campaigns. Each campaign was planned after evaluating previous results to adequately address local geogenic and anthropogenic complexities. A geological map at a scale of 1:50,000 (GeoČR50) and a map of mining sites, both provided by the Czech Geological Survey (www.geology.cz),

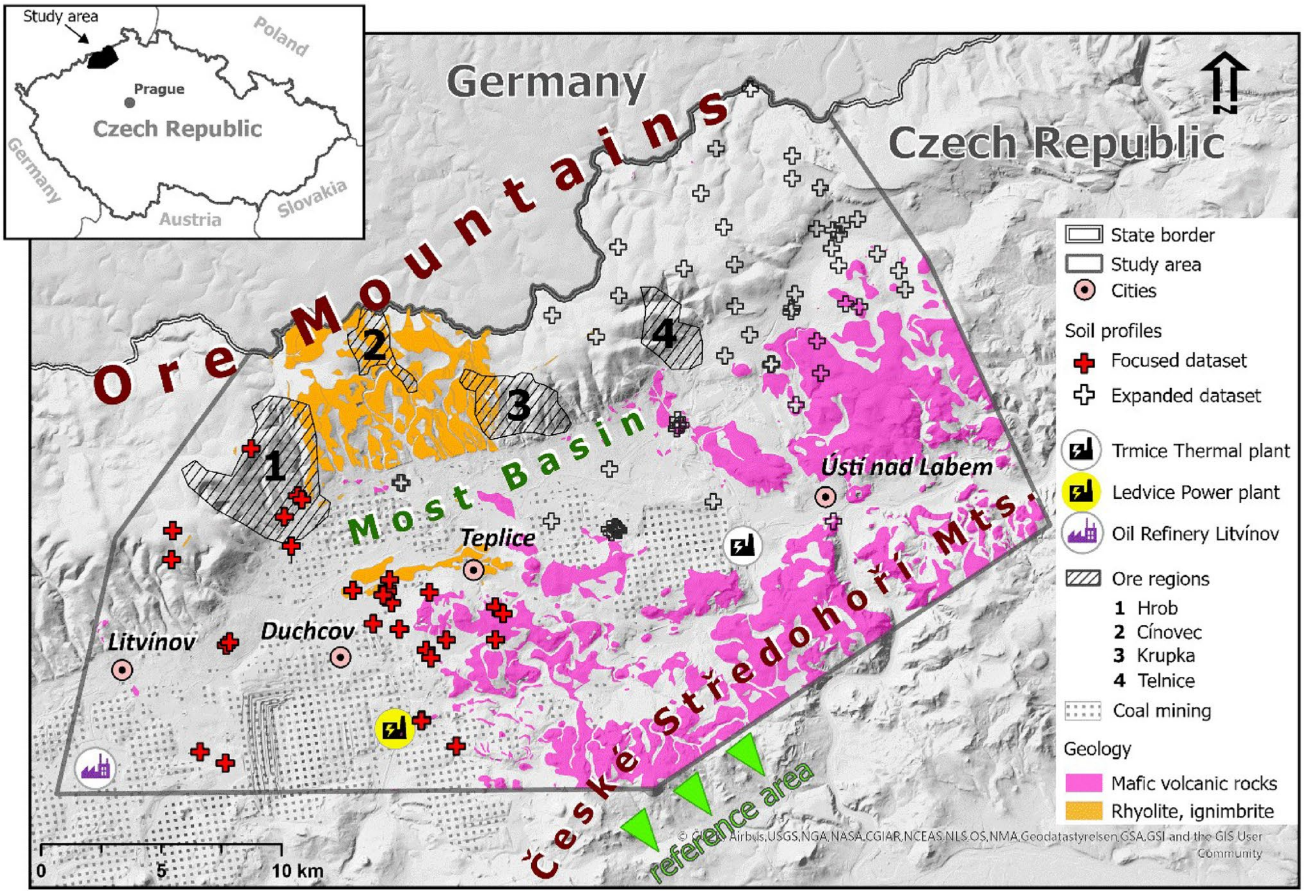

Fig. 1 Study area showing the distribution of major classes of anomalous rocks, ore regions, the location of major coal combustion plants, and the positions of sites, where soil depth profiles were sampled. The composition of soils in the reference area (green triangles) was taken from Adamec et al. (2024)

were used in planning the sampling sites and interpreting the data. Two sets of samples and data are distinguished: all samples, of which position is shown in Fig. 1, are denoted expanded dataset, while samples located in the southwest part of the study area, where most coal-related contamination sources have been situated, are denoted focused dataset.

Expanded dataset contained 175 samples from 34 soil profiles along and outside the extent of the Most Basin floor. They covered all bedrock geologies met in the focused dataset. The dataset was obtained newly for this work. Focused dataset included 152 samples from 29 soil profiles, of which 45 samples from 10 profiles were sampled newly and geogenic anomalies, while the rest of samples were taken from preceding work (Adamec et al. 2024). Sampling targeted agricultural and forest soils (25 samples), positioned at least 20 m from roads and 100 m from buildings to minimize the impact of non-coal contamination. Soil profiles were obtained using a soil auger corer (Eijkelkamp, the Netherlands), preferably to a depth of approximately 1 m, if the soil profiles allowed, with sampling steps of 10–20 cm.

The soils were sieved with a 2 mm mesh and dried under laboratory conditions. In focused dataset, the pseudo-total content analysis was performed to determine As, Be, Cd, Pb, and Sb. For this, soil samples (<2 mm, 0.5 g) were extracted using 6 ml HCl and 2 ml HNO_3 in a Multiwave 5000 microwave (Anton Paar) following the European standard “EN ISO 54321:2021: Soil, treated biowaste, sludge and waste—Digestion of aqua regia soluble fractions of elements.” After centrifugation, the supernatant was diluted and analysed using an Agilent 7900 ICP-MS (Adamec et al. 2024).

A portion of all samples (expanded dataset) was pulverized in a planetary micromill and analysed by X-ray fluorescence spectrometry (XRF) to obtain total concentrations of As, Cu, Pb, and Zn as risk elements, as well as Al, Si, K, Ca, Ti, Mn, Fe, Rb, Sr, and Zr as lithogenic elements. Pulverized samples were poured into conventional nylon XRF cells with Mylar foil bottoms. Analyses were performed using an Epsilon 3X spectrometer (PANalytical, the Netherlands) equipped with an X-ray tube (Ag cathode, up to 50 kV) and a Peltier-cooled large-area silicon-drift detector.

The quantification of element concentrations was performed with single-element calibration using a set of certified reference materials, as described previously (Adamec et al. 2024; Hošek et al. 2024).

Subsamples for polycyclic aromatic hydrocarbons (PAHs) were sieved (<2 mm) in the field, packed in aluminium foil, and freeze-dried in the laboratory. PAHs were determined after ethyl acetate extraction following the QuEChERS procedure. Analysis was performed using a 7890B gas chromatograph with a 7000D GC/MS detector (Agilent Technologies, USA).

Granulometric curves were obtained using a Bettersizer S3 Plus with the Bettersize Laser Particle Size Analysis System V8.10 (Bettersize, China).

Data processing

The major rock types on the geological map were selected from geological maps at a scale of 1:50,000 (GEOČR50). Ore regions were manually delineated according to the map of mine workings provided by the Czech Geological Survey, which was accessed as a map service in ArcGIS (https://mapy.geology.cz/arcgis/rest/services/Dulni_Dila/dulni_dila/MapServer). The fifth-generation digital elevation model of the Czech Republic (DMR 5G) from the Czech Office for Surveying, Mapping, and Cadastre (ČÚZK) was created in 2011 and 2018 and was used in Fig. 1. This model was included in the project as a raster from Image Service. The data were processed using ArcGIS Pro 3.2.0 software (Esri, USA).

Robust algorithms were favoured for data mining due to their lesser sensitivity to outliers compared to ordinary least squares methods (Bábek et al. 2015; Matys Grygar and Popelka 2016). Multivariate linear robust regression (MLRR) with the MM-estimator was employed to simultaneously predict the relationship between a single dependent variable (risk element) and multiple independent variables (lithogenic elements) in soils (Filzmoser and Nordhausen 2020). MLRR was performed stepwise, beginning with all lithogenic elements as explanatory variables and systematically removing those that did not pass the t-test of statistical significance at $P = 99\%$.

Principal component analysis (PCA) was conducted to investigate the geochemical characteristics of the soils and the relationships between the studied elements. PCA is a widely used dimensionality reduction technique that constructs a new set of uncorrelated and orthogonal variables, often serving as a preliminary step in analysis of high-dimensional data (Jolliffe, 1986). However, PCA is very sensitive to outliers, which are common in geochemical data and can lead to misleading conclusions. To address this issue, the robust version of principal component analysis (RPCA) was employed. The core of RPCA is the minimum covariance determinant (MCD) method, an extremely robust estimator of multivariate location and scatter. Details of this method can be found in Hubert and Engelen (2004) and Hubert et al. (2018). Element concentrations were logarithmically transformed prior to PCA.

Statistical analyses were performed using the R statistical platform (R Core Team 2023). The original granulometric curves were smoothed and transformed using the centred logratio (clr) transformation, a common preprocessing step in Bayes space methodology (van den Boogaart et al. 2014). This enabled the statistical analysis of the contained relative information via standard functional data analysis (Ramsay and Silverman 2005). The individual clr-transformed granulometric curves were plotted as a heatmap complemented by a clustering dendrogram. The dendrogram utilized the complete linkage method (e.g., Filzmoser and Varmuza 2009), where hierarchical clustering was performed based on the smallest Euclidean distance between pairs of observations.

Statistical analyses were performed using the R statistical platform (R Core Team 2023).

Results

Identification of Geogenic anomalies

Local geogenic anomalies have been identified through soil depth profiles, where risk elements achieved high concentrations relative to common standards throughout the entire soil profiles (Sect. "Study area" in Supplementary Information, Fig. S3 to S7). The overview of anomalous rocks is presented in Table 1, along with bedrock information from GeoČR50, references to studies supporting the anomaly assignment, and typical characteristics of the soils, which are described in more detail below in the text. Soil anomalies are not limited to the mapped extent of these anomalous rocks, as erosion has spread their weathering products to surrounding soils through aeolian or slope transport in the Pleistocene, and possibly also through anthropogenic activities in the late Holocene. On the basin floor, the relevant areas include (i) sites downslope of the mineralized regions in the Ore Mountains (e.g., Jeníkov and Lom) and (ii) sites around flat elevations formed by effusive volcanic rocks in the basin floor (e.g., Hudcov).

Table 2 shows that topsoils have higher concentrations than subsoils for several risk elements. Notably, the largest median concentrations for Cd, Pb, Sb, and Zn, along with elevated As levels, were found in the soils of the Ore Mountains and those overlying felsic igneous rocks, regardless of sampling depth (Table 2). The highest median concentration

Table 1 Soil profiles displaying concentrations of risk elements that are naturally elevated throughout the entire depths and attributed to bedrock geology, including mineralization near ore regions

Sites	Elements	Interpretation	Typical signs
Hudcov, Hrob (Fig. S5, S6)	As, Be	Geogenic anomaly related to felsic effusive rocks (Cicchella et al. 2015; Armiento et al. 2022; Murray et al. 2023)	Geological map, elevated K/Al; sandy-silty soils
Jeníkov	As, Pb	Downslope from ore region 2 and/or felsic effusive rocks, shown in Fig. 1	Elevated Rb/K, sandy-silty soils
Sites near ore regions 1 and 4 (Fig. S3, S4)	As, Pb, Sb, Zn	Local mineralization (Ore Mountains) (Matys Grygar et al. 2016; Hošek et al. 2024)	Geological map; As ~ Pb; elevated Rb/K, elevated K/Al; sandy-silty soils
Soils on mafic effusive rocks (Fig. S7)	Cu	Geogenic anomaly related to basalt bedrock (Ballabio et al. 2018; Matys Grygar et al. 2023; Zhang et al. 2020a, b; Adamec et al. 2024)	Geological map; elevated Fe and Ti; clayey-silty soils

Table 2 Medians of pseudototal (subscript AR) and total (subscript TOT) concentrations of the risk elements in the focused dataset, depending on bedrock geology and sampled soil depths. Bold fonts highlight extreme element concentrations in the locally anomalous bedrock. Reference concentrations are also listed

Soil parent rock (number of samples	Depth (cm)	As_{TOT} (mg kg^{-1})	As_{AR}	Be_{AR}	Cd_{AR}	Pb_{TOT}	Pb_{AR}	Sb_{AR}	Zn_{TOT}	Zn_{AR}
Cretaceous (26)	0–30	21	37	1.94	0.39	52	30	0.74	90	81
	> 30	16	12	**1.30**	0.18	33	15	0.30	53	44
Felsic effusive (20)	0–30	53	43	1.85	0.33	53	47	1.55	90	94
	> 30	32	**145**	**1.77**	0.12	47	26	**1.83**	105	65
Mafic effusive (18)	0–30	18	21	2.02	0.48	52	28	0.90	108	111
	> 30	13	11	0.96	0.18	43	9	0.36	89	86
Miocene sediments (20)	0–30	27	17	1.67	0.30	56	31	0.77	114	64
	> 30	32	12	0.90	0.06	55	18	0.71	112	50
Sediments Pleistocene (47)	0–30	22	19	1.77	0.18	51	30	0.84	110	86
	> 30	20	11	**1.43**	0.21	29	19	0.53	75	49
Ore Mts. (21)	All	42	**91**	1.01	**0.42**	**64**	**58**	**1.79**	127	**121**
Reference area outside the basin[1]	0–30	32	10	1.34	0.24	32	24	0.54	88	74
	> 30	15	6	1.13	0.12	15	15	0.30	45	47
Czech Republic[2]	0–40		9	0.94	0.20		22			66
EU topsoils[3]			8	0.68	0.22		21	0.35		53

[1] Adamec et al. (2024) dataset, agricultural area southeast of the Most Basin (its position is indicated by green arrows in Fig. 1)

[2] State database of the risk element concentrations in agricultural soils determined by AR (Matys Grygar et al. 2023)

[3] European topsoils from ploughed horizons Ap in the south Europe subset from Reimann et al. (2018)

of Be and elevated As were found in soils overlying felsic effusive volcanics. Soils on these rocks exhibit considerably higher concentrations of the risk elements than the medians in soils from the reference locality southeast of the Most Basin (green arrows in Fig. 1; more details in Adamec et al. 2024) and elsewhere in Europe.

The major features of bedrock-related element variability in the studied soils are visualized in the PCA plots (Figs. 2 and 3). The local geogenic anomalies (Table 1) represent the end members of the basin floor soils. PCA was performed on the total element content in the expanded dataset (Fig. 2) as well as on a combination of total and pseudototal concentrations in the focused dataset (Fig. 3). While the results are principally similar, the element loading vectors have different orientations in both versions, indicating that the principal components have exchanged meanings. Low percentages of explained element concentration variability in both PCAs documents non-linear interelement relationships and local variability of soils. This prevents reliable use of geochemical indices in the study area, but still PCA as exploratory tool makes it possible to decipher major geogenic controls of risk element concentrations.

The bedrock end members identified include felsic effusive rocks and rocks from the Ore Mountains, which show high concentrations of K, Rb, and Si, as well as As, Pb, and Sb. Cretaceous sediments, particularly in deeper horizons, exhibit high Ca and low concentrations of risk elements, while mafic volcanic rocks are characterized by high levels of Cu, Fe, Mn, and Ti.

In Fig. 2, PC1 illustrates the proportions of mafic and felsic end members, while PC2 reflects the contribution of the Cretaceous end member with low Zn and other risk elements relative to the other end members. Conversely, in Fig. 3, PC2 represents the mafic and felsic end members,

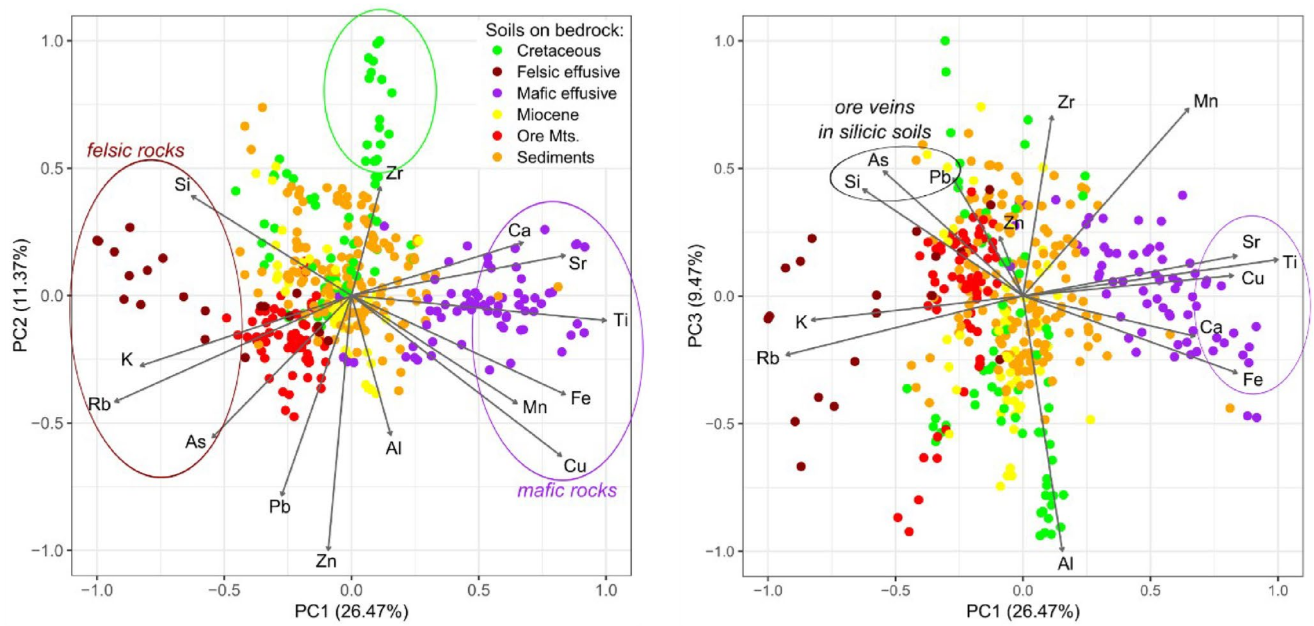

Fig. 2 Principal Component Analysis (PCA) of total concentrations in soils from the expanded dataset. The ellipses indicate lithogenic end members on the PC2 vs. PC1 plot, while the PC3 vs. PC1 plot illustrates ore-vein associations

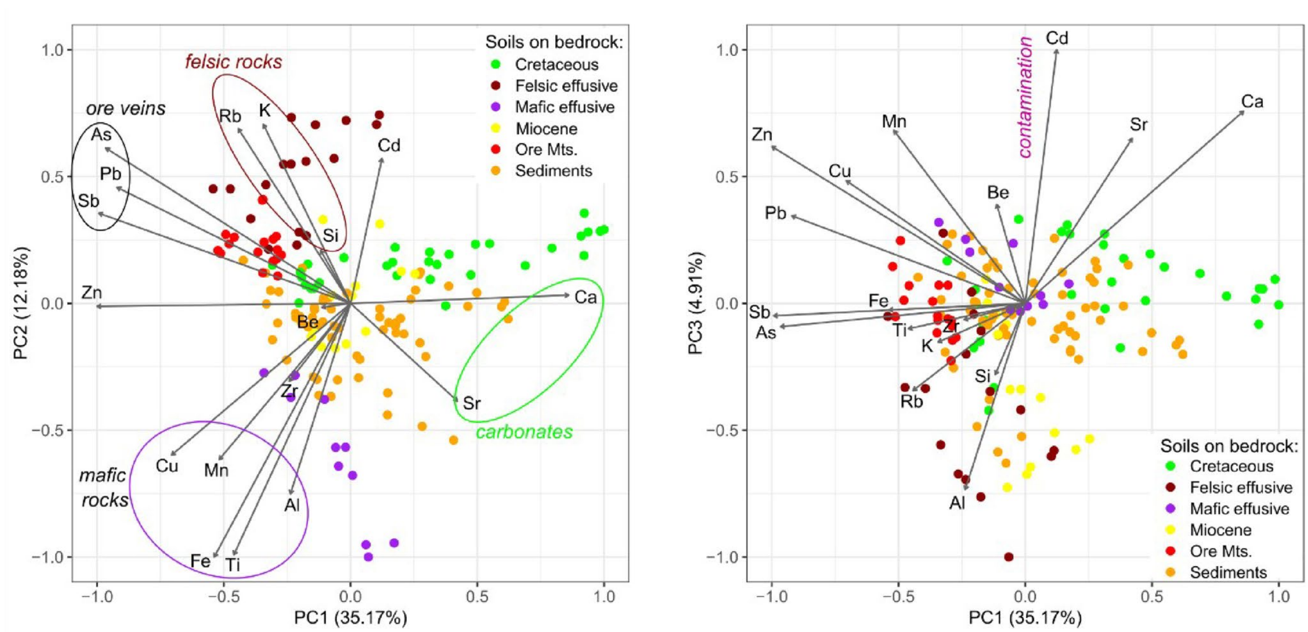

Fig. 3 Principal Component Analysis (PCA) of pseudototal concentrations (As, Be, Cd, Cu, Pb, Sb, Zn) and total concentrations (other elements) in soils from the focused dataset. Ellipses indicate lithogenic end members (PC2 vs. PC1 plot), ore-vein associations (PC2 vs. PC1 plot), and Cd-dominated contamination (PC3 vs. PC1 plot)

and PC1 indicates the contribution of the Cretaceous end member.

PC3 in both figures may represent contamination signals; however, this contamination likely originates from both ore veins in the geogenic anomalies and anthropogenic activities. Notably, there is a close association of As, Pb, Sb, and commonly nearby vectors of Zn, which is typical for polymetallic sulphide ores (Zhang et al. 2020a, b) and may reflect the local geogenic anomalies (Table 1). Elements such as Be, Cd, and Zn generally exhibit weak lithogenic associations and behave independently of each other, indicating that their concentrations are not mainly controlled by a common source.

Table 3 MLRR coefficients (intercepts and slopes) for total risk element concentrations in the expanded dataset (175 samples from 34 soil profiles), along with UCC contents of the risk elements (in square brackets) and UCC element ratios. Bold font indicates MLRR slopes that exceed the UCC ratios

Variable	As (mg kg^{-1})		Cu (mg kg^{-1})		Pb (mg kg^{-1})		Zn (mg kg^{-1})	
	Intercept	UCC[1]	Intercept	UCC	Intercept	UCC	Intercept	UCC
	-0.9	[4.8]	4	[28]	12	[17]	-19	[67]
	Slopes	Ratios	Slopes	Ratios	Slopes	Ratios	Slopes	Ratios
Al (%)					-1.7	2.1	7	8.2
Si (%)	**0.58**	0.15			0.4	0.55		
K (%)					**11**	7.3	18	29
Ti (%)	-8.8	13	5.1	73	-10	44		
Fe (%)	**1.5**	1.2	4.0	7.1	**8**	4.4	6	17
Zr (mg kg^{-1})	0.012	0.025						
Adj. R^2	0.26		0.78		0.73		0.53	

Table 4 MLRR coefficients for pseudototal concentrations of risk elements in the focused dataset (152 samples from 29 profiles), alongside the median European pseudototal soil concentrations of these risk elements (in square brackets) and the ratios of European pseudototal risk element concentrations to total lithogenic element concentrations. Note that cd did not yield relevant fits

Variable	As (mg kg^{-1})		Be (mg kg^{-1})		Pb (mg kg^{-1})		Sb (mg kg^{-1})		Zn (mg kg^{-1})	
	Int.	EU[1]	Int.	EU	Int.	EU	Int.	EU	Int.	EU
	-8	[8]	0.5	[0.68]	11	[21]	-0.2	[0.35]	-26	[53]
	Slopes	Ratios	Slopes	Ratios	Slopes	Ratios	Slopes	Ratios	Slopes	Ratios
Si (%)							**0.014**	0.011		
K (%)	**18**	5.4	**0.6**	0.46					**61**	35
Ti (%)									62	133
Mn (mg kg^{-1})			**0.005**	0.001						
Fe (%)	2.1	2.9								
As (mg kg^{-1})					0.18	2.1	0.005	0.035	-0.10	5.3
Rb (mg kg^{-1})	-0.06	0.1	-0.003	0.0085	0.04	0.26	0.002	0.004	-0.20	0.66
Zr (mg kg^{-1})	-0.03	0.032							-0.09	0.21
Adj. R^2	0.36		0.21		0.50		0.91		0.53	

[1] Reimann et al. (2018) for pseudototal concentrations and Reimann et al. (2012) for total concentrations in south European soils

Figures 2 and 3 illustrate that basin floor cover classified as Pleistocene sediments in the GeoČR50 exhibits geochemical variability comparable to that of their parent rock end members. The points labelled as sediments in both figures occupy the entire area between the mafic, felsic, and Cretaceous end members, suggesting that the influence of the geogenic anomalies related to felsic rocks and ore veins can also be expected in these sediments. The Pleistocene deposits on the Most Basin floor are categorized as colluvial, alluvial, or aeolian sediments in the GeoČR50, without any indication of their lithogenic origin.

This mixing of depositional mechanisms implies that variable percentages of the major end members can be expected, but such distinctions cannot be extracted from the geological map. This limitation can potentially be addressed using geochemical tools: the signatures of lithogenic elements can help decipher the soil lithogenic origin (as shown in Table 1; Figs. 2 and 3). The lithogenic element(s) could theoretically be tested as predictors for natural (local background) concentrations of the risk elements in soils; Multivariate Linear Robust Regression (MLRR) was employed for this purpose to identify a background function for the natural concentrations of risk elements.

Use of interelement relationships in soils

Multivariate Linear Robust Regression (MLRR) was applied to both datasets, excluding samples from the top 30 cm, where significant anthropogenic contamination is expected. The resulting parameters of the local background curves are presented in Tables 3 and 4. These regressions yielded good results (R^2 >0.7) for predicting the natural concentrations of Cu, total Pb, and Sb. Modest but still plausible results (R^2 approximately 0.5) were obtained for Zn and pseudototal Pb.

In all cases, the results of MLRR (Tables 3 and 4) can help elucidate the basic factors controlling the distribution of risk elements in the Most Basin floor by comparing the regression slopes with upper continental crust (UCC) element ratios. The regression slopes for Cu, Pb, and Zn are usually similar as the global reference, represented by the UCC element ratios. Cu and Zn exhibit significant positive slopes in relation to Al, Fe, and/or Ti, which are typically associated with finer size fractions (Matys Grygar and Popelka 2016) and mafic lithogenic origin. In contrast, elements with positive slopes related to Si as an explanatory variable, comparable to or exceeding the UCC element ratios (such as total As and Pb, as shown in Table 3, and pseudototal Sb, as shown in Table 4), are typically more abundant in coarser (silicic) soil fractions or soils with a greater felsic contribution. This suggests they are more likely to be present in primary coarse-grained ore minerals or influenced by geological origin.

The prevailing lithogenic control of As and Pb in the Most Basin soils is evidenced by their bedrock-driven association with K and Rb, as identified by PCA (Fig. 2) and by correlation with K and Si, as identified by MLRR (Tables 3 and 4). However, the MLRR background functions did not quantitatively remove the effects of anomalous soils impacted by ore veins in the Ore Mountains. PCA demonstrated that the felsic effusive rocks and geogenic anomalies from the Ore Mountains are characterized by elevated K concentrations, to which elevated Pb is also associated (Fig. 2). Elevated K or Rb/K can serve as diagnostic indicators of the geogenic control of risk elements in basin floor soils (Table 1, last column).

The red rectangle in Figs. 4A and B illustrates that the majority of soil samples with anomalously high Pb also exhibit anomalously high K/g(Al, Si), where g(Al, Si) represents the geometric mean of Al and Si concentrations. However, the MLRR-corrected Pb concentration, denoted as Pb* in Fig. 4, is not simply proportional to the lithogenic anomaly of K/g(Al, Si), indicating that MLRR cannot provide a quantitative correction for the anomalies related to the Ore Mountains and felsic rocks.

The behaviour of As (Fig. 4C and D) and Zn, which is also elevated in some Ore Mountains soils, was similar to

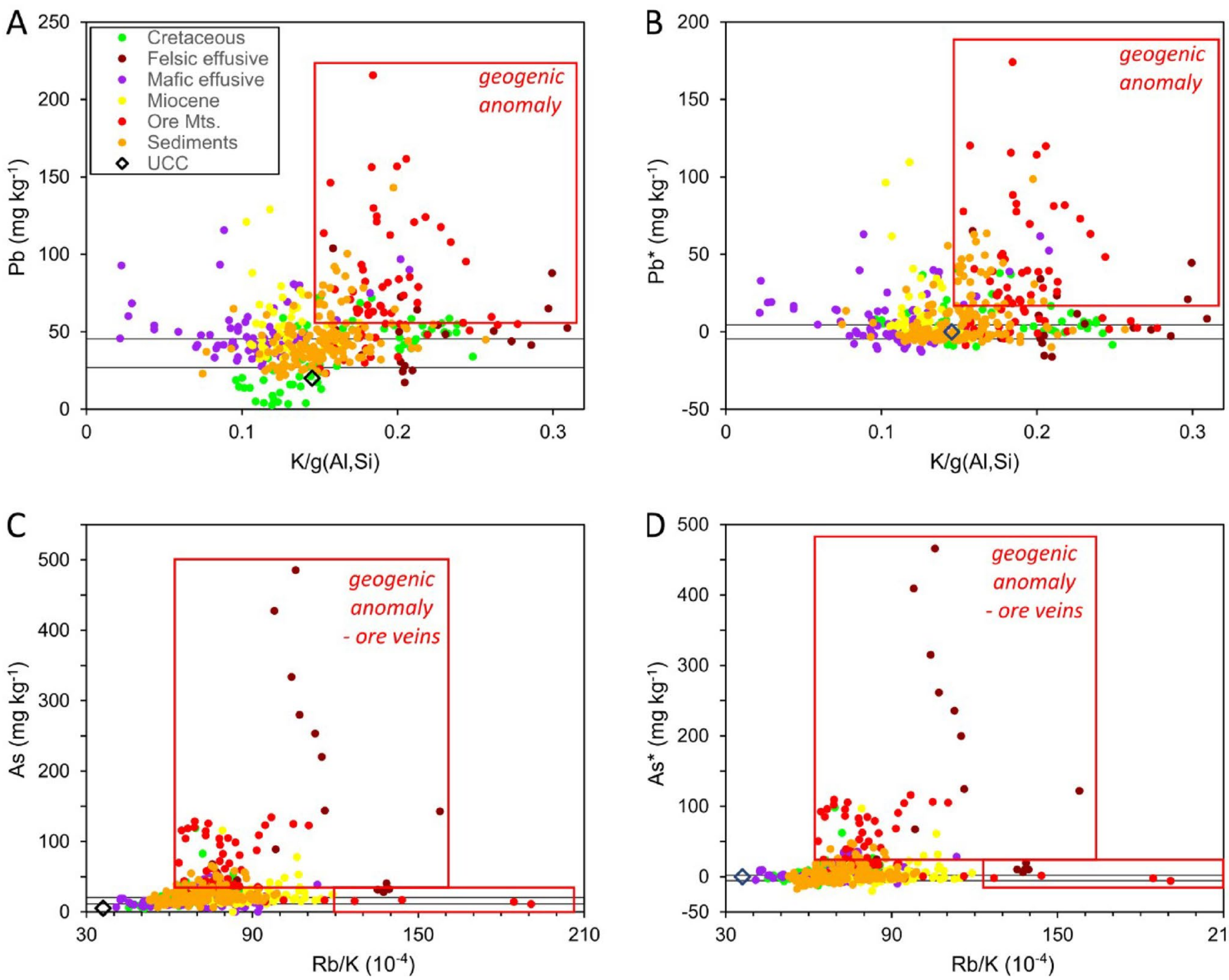

Fig. 4 Concentrations of total Pb (**A**), total Pb corrected using the Multivariate Linear Robust Regression (MLRR) equation (**B**), total As (**C**), and total As corrected using the MLRR equation (**D**) plotted against normalized K (**A**, **B**) and the rubidium/potassium (Rb/K) ratio (**C**, **D**) for the expanded dataset. The black rectangle marks the 2nd to 5th deciles of results, while the red rectangles highlight contamination associated with anomalous K/g(Al, Si) or Rb/K relative to the upper continental crust (UCC) composition

Pb. The relatively small regression coefficients for As, Be, and Pb does not imply that the MLRR approach failed completely; the lithogenic element concentrations successfully corrected the soil risk element contents for the lithological variability of soils, as evidenced by the narrowing of the risk element concentration distribution in non-contaminated samples. This narrowing is highlighted by the black rectangles in Figs. 4A and B, which depict the width of the 2nd to 5th deciles of Pb or Pb* concentrations in deeper soil horizons with autochthonous uncontaminated soils. The width of these intervals is much smaller for Pb* (9 mg kg^{-1}) than for raw Pb (19 mg kg^{-1}).

Granulometric curves and soil lithogenic origin

Certain origin of soils in the Most Basin can be distinguished based on granulometry, as indicated in the last column of Table 1. The granulometric curves of the focused dataset have been processed using Bayes space methodology and subsequently clustered (Fig. 5), as described in the methodological section. A substantial sand component > 100 μm is always present, while the finest clay < 1 μm is absent in soils derived from the felsic rocks of the Ore Mountains (cluster B in Fig. 5). In contrast, sand > 100 μm is absent, and the clay component is significant in soils derived from Cretaceous marls (with the finest clay, cluster F in Fig. 5) and basalts (mainly clay and silt, clusters D and F in Fig. 5).

Concentration of Ca, which is elevated in soils derived from Cretaceous sediments (mostly marls with considerable amount of carbonates), is also useful for distinguishing soil lithogenic origin. True loess consists mostly of sorted silt with possible coarse clay particles formed by later pedogenesis (cluster D in Fig. 5, high densities in grain sizes between ca. 1 and 60 μm). However, the considerable sand admixture to silt in samples classified as loess in GeoČR50 (cluster A in Fig. 5) suggests that these soils have received significant alluvial components from the slopes of the Ore Mountains. This comprehensive examination of the soils indicates that the lithogenic origin of some soil samples is not always accurately classified based on GeoČR50 (Table 5).

Table 5 indicates that approximately one-third of the topsoil samples assigned as the Cretaceous sediments according to GeoČR50, are actually Pleistocene sediments with sand component > 100 μm, primarily located in cluster C of Fig. 5, which have very similar grain/size distribution as the Ore Mountains soils in cluster B. The most challenging scenario for studies on anthropogenic contamination arises from changes in soil lithogenic origin within single soil depth profiles, in which soils developed from the underlying Cretaceous sediments are overlain by topsoil derived from Pleistocene sediments with a considerable sand fraction. This sand could only originate from the Ore Mountains or felsic effusive rocks, because neither mafic effusive rocks nor local Cretacous sediments do neither contain quartz sand nor cannot produce sand by their weathering. The change of lithogenic origin must inevitably be associated with geogenically elevated concentrations of several risk elements, which could possibly be misinterpreted as

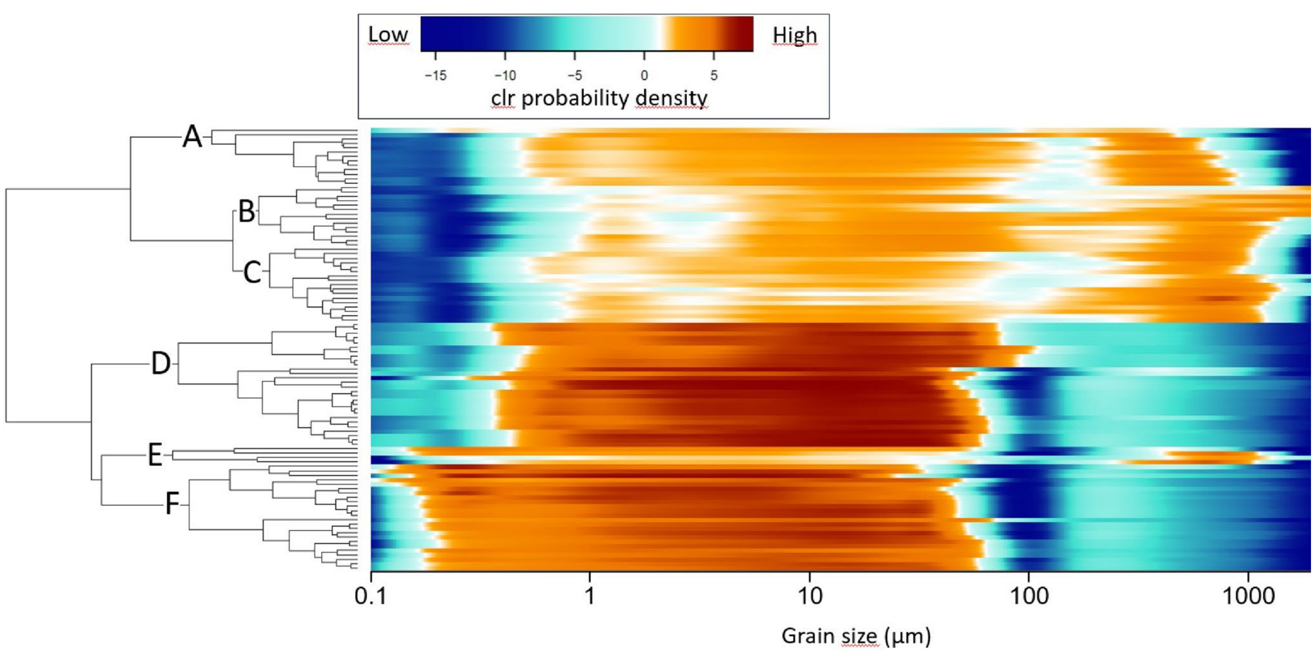

Fig. 5 Clustering of soil samples from the focused dataset according to the probability densities calculated from granulometric curves. Warm (reddish) hues indicate a dominance of the corresponding grain sizes, while blue hues signify their absence. Clusters of similar grain-size distributions (A to F) are discussed in the text

Table 5 Misfits between soil lithogenic origin classification and information provided by GeoČR50. The cluster labels refer to Fig. 5

Assignment according to GeoČR50	Decisive features: grain size [cluster label]; geochemistry	GeoČR50 assignment confirmed	GeoČR50 assignment not confirmed	Characteristics of the poorly assigned (allochthonous) soil samples [cluster label]
Loess	Sorted silt dominant (>50%), sand absent [D]; elevated Zr content[1] (>300 mg kg^{-1})	21 samples	21 samples	Sorted silt<40%, medium and coarse sand (>0.2 mm)>20% [A or C]
Cretaceous	Dominant clay size fraction [F or E]; elevated Ca content>10%	9 samples	20 samples	Negligible clay content, prevailing sand and silt [B or C]; Ca <1%; in some samples Zr >300 mg kg^{-1}

[1] criterion used by Waroszewski et al. (2017)

surface contamination from emissions if soil origin was not considered. This pattern of soils derived from Cretaceous sediments and covered by allochthonous sandy deposits was observed in a total of four soil profiles.

In half of the soil samples classified as residing on loess sediments according to GeoČR50, the sand content proved to have an alluvial origin (Table 5). These poorly classified soils are located downslope from the Ore Mountains and exhibit very similar grain-size features (clusters A and C) to the soils from the Ore Mountains (cluster B). In these misclassified soils, geogenic enrichment of As, Be, Pb, and Sb can be expected (Tables 1 and 2), which could be mistakenly interpreted as anthropogenic contamination.

Topsoil enrichment of risk elements and PAHs in the ledvice power plant area

The impact of coal combustion can be assessed through PAH contamination in topsoils, independent of soil lithogenic origin and geochemical features. In Fig. 6, topsoil PAH and some risk element concentrations are plotted against the distance from the Ledvice Power Plant, which is situated in the centre of the study area (Fig. 1) in the Most Basin. Notably, topsoil concentrations of Cd, Zn, and PAHs are particularly high at a site Ledvice, only 1 km from the power plant. Maximal PAH concentrations were found in Lom, a forested site; tree crowns are known to scavenge atmospheric (dust) contamination more efficiently than agricultural soils.

Topsoil concentrations of As, Sb, and Pb are highest near the geogenic anomalies (Tables 1 and 2) in the Ore Mountains, around ore regions, and in the basin floor close to these areas (Fig. 6; the position of individual sites is shown in Fig. 1). There is no simple decline of the soil risk element concentrations with growing distance from the Ledvice power plant, documenting major impact of site-specific soil chemistry in the target area. Unfortunately, the impact of geogenic anomalies and ore regions cannot be quantified or subtracted from the actual concentrations of lithogenic elements using MLRR, that result in uncertainty in unequivocal contamination assignment around the Ledvice Power Plant to the anthropogenic impact of the coal combustion. This uncertainty prompts for using median statistics for the data mining rather than evaluation of individual soil samples with locally maximal contamination.

The median soil concentration was calculated for samples within a 10 km radius of the Ledvice Power Plant (excluding sites located directly on the recognised geogenic anomalies), as indicated by the red line in Fig. 6. The contribution of the power plant can be assessed by comparing this median to the median concentration from the reference area outside the Most Basin behind the České Středohoří Mountains (green arrows in Fig. 1). The results of those comparisons are in Table 6.

Discussion

Use of lithogenic elements to decipher the origin of elevated risk element concentrations

Data treatment in elemental contamination studies should effectively distinguish anthropogenic contributions to risk elements in environmental samples. This task could be considerably facilitated by correcting for natural compositional variability using lithogenic element concentrations, which enhances the sensitivity of detecting contaminated samples as outliers, as it was successfully performed with sediments or soils derived from sediments (Sterckeman et al. 2006; Bábek et al. 2015; Lučić et al. 2023). Adamec et al. (2024) did try this approach for the Most Basin soils, but the attempt was not successful. In this paper, novel sampling and MLRR were employed once more, utilizing various lithogenic element concentrations as predictors for risk element concentrations (Tables 3 and 4). This novel attempt respected all rational recommendations to separate geogenic and anthropogenic contributions to soil risk element concentrations. The regression analysis only included statistically significant explanatory variables to prevent data overfitting, thereby keeping the regression coefficients interpretable. Overfitting, or the use of overly complex statistical methods, poses a significant challenge to interpretation (Greenacre 2019).

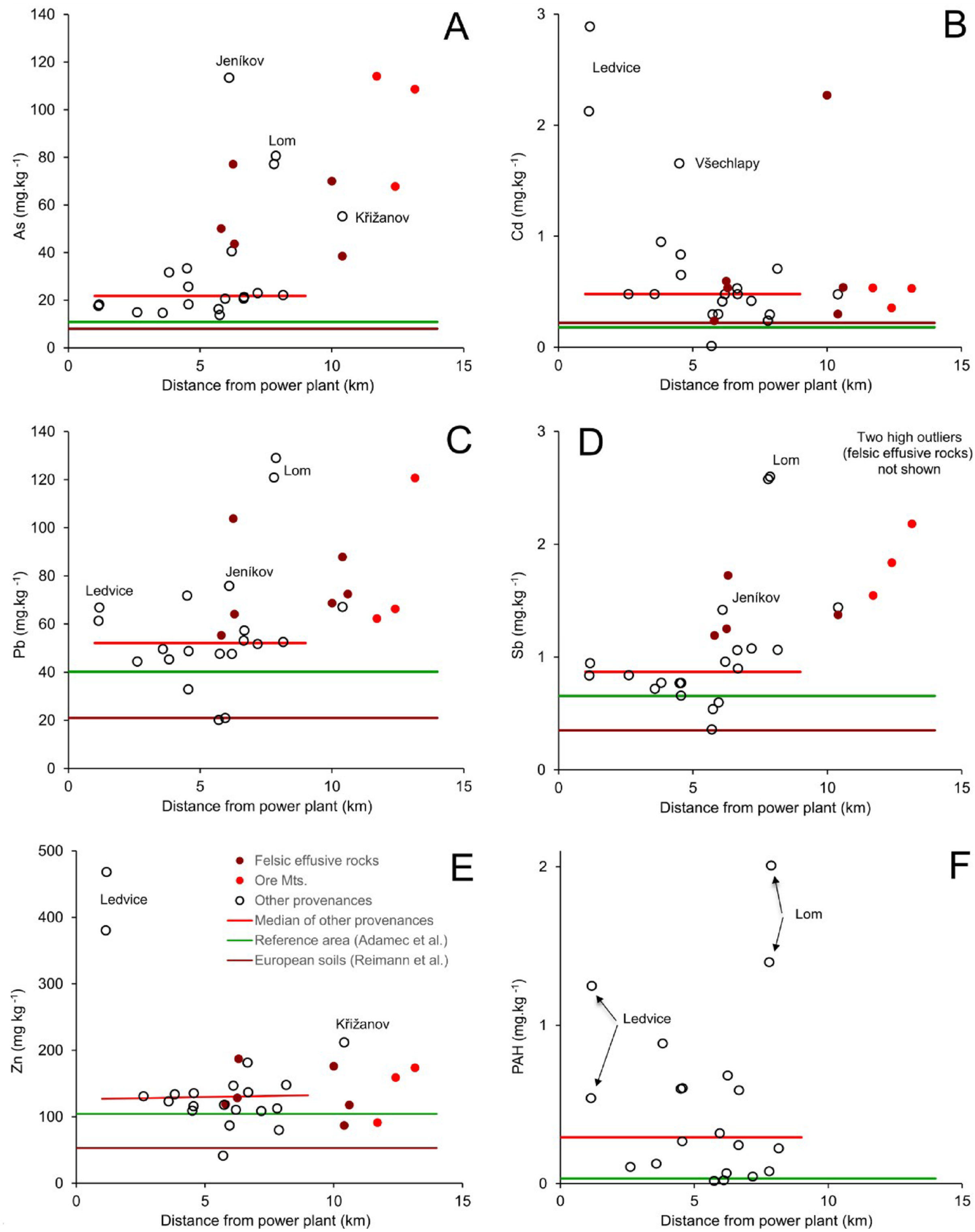

Fig. 6 Contaminant concentrations plotted against the distance to the Ledvice Power Plant. Also shown for comparison are results from the reference area (Adamec et al. 2024) and EU soils (Reimann et al. 2018)

Table 6 Medians of the basin floor topsoil concentrations of risk elements and PAH sums, along with differences between these values and the median in the reference dataset. Also listed are median contents in EU soils (Reimann et al. 2018) and the preventive limit for Czech agricultural soils

Contaminant	Basin floor near Ledvice	Difference[1]	EU soils	Preventive limit[2]
	(mg kg^{-1})	(mg kg^{-1})	(mg kg^{-1})	(mg kg^{-1})
As	22	11	8	20
Be	1.9	0.7	0.7	2
Cd	0.48	0.30	0.22	0.5
Cu	30	5	19	60
Pb	52	12	21	60
Sb	0.87	0.21	0.35	-
Zn	104	28	53	120
ΣPAH	0.29	0.26	-	1

[1]Difference of medians for the basin floor around Ledvice and median in the reference area separated from the basin floor by the České Středohoří Mountains (Adamec et al. 2024)

[2] Czech legislative for agricultural soils, 153/2016

For instance, Sun et al. (2024a, b) employed partial least squares regression (PLSR) and included up to ten explanatory variables for each risk element in their study, although meaningful regression coefficient increases were primarily observed with approximately five variables. Notably, a considerable proportion of their regression coefficients were negative, making them difficult to interpret in physical or chemical terms. Negative slopes may compensate for the non-linearity of interelement relationships (Matys Grygar and Popelka 2016; Jung et al. 2016). Sun et al. (2024a, b) did not address these slopes from a geochemical perspective, a focus of our investigation.

The association of risk elements with certain lithogenic elements can arise from shared influences such as grain size control (Matys Grygar and Popelka 2016; Ballabio et al. 2018), dilution by organic matter, quartz, or calcium carbonates (McKinley et al. 2016; Lučić et al. 2023), common geological origins involving primary minerals, or shared fates during pedogenesis (involving secondary minerals formed by bedrock weathering or sorption of risk elements onto lithogenic minerals like Fe or Mn oxides, as noted by Matys Grygar and Popelka 2016). This multitude of potential controlling mechanisms explains why the set of lithogenic element predictors cannot be predetermined and must be tested and tailored for each specific element and study area, as done in this work.

The MLRR and PCA were partly successful and showed several important geochemical factors controlling risk element concentrations, in particular related to lithogenic origin of soils. Potassium concentration is less commonly used as a lithogenic element predictor (Tables 3 and 4); however, in the Most Basin it clearly reflects the association of several risk elements with felsic effusive rocks and the bedrock of the Ore Mountains, that is also composed of felsic rocks (Table 1; Figs. 2 and 3). MLRR successfully decreased variability in uncontaminated soil samples, as demonstrated by the narrowing of the 2nd to 5th deciles of results in Fig. 4. However, MLRR did not completely eliminate anomalies related to the Ore Mountains and felsic volcanic rocks, revealed in this study (Table 1). The relationship between extremely high concentrations of As, Pb, Sb, and Zn, and anomalously high K or Rb/K is probabilistic, indicating a higher likelihood of association rather than being deterministic, meaning it is not strictly controlled arithmetically; thus, such elements can also be absent (Fig. 4).

The association of the risk elements As, Pb, and Sb with K and Rb, including their extreme concentrations, supports the prevailing geogenic control of these risk elements in the soils of the Most Basin. Mineralization in the Ore Mountains has been linked to intrusions of felsic volcanic rocks, particularly granites, which are generally rich in K and Rb (a common characteristic of such rocks, Zhang et al. 2020a, b). Elevated Rb/K ratios have been observed in the granitic and mineralized sub-catchment of the Ohře River in the Ore Mountains (Matys Grygar et al. 2016). An increased Rb/K ratio is often used as a measure of the degree of magma differentiation during granite formation and metamorphosis, which includes the formation of Sn ore veins in the Ore Mountains and elsewhere (Breiter et al. 2017; Sun et al. 2024a, b). However, it is not possible to simply predict As and Pb concentrations in geogenically anomalous soils using Rb, K, or Rb/K ratios using some geochemical index such as enrichment factor or establishing some geochemical background value for soils on specific rocks.

In the dataset examined in this study, the majority of soils with extremely high levels of As also exhibited elevated Rb/K ratios, although this relationship is not deterministic (Fig. 4C and D). Both lithogenic proxies have also further limitations: Stafilov et al. (2018) noted that K concentrations in topsoils can be increased by fertilizers. Additionally, Rb/K ratios can be elevated due to intensive pedogenesis, as Rb is less geochemically soluble in exogenic environments than K (Rudnick and Gao 2003). Consequently, elevated Rb/K ratios can also be observed in Cretaceous and Miocene sediments (Fig. 4C and D), where neither As nor Pb levels are elevated. Therefore, the analysis of lithogenic element concentrations in soils should also be employed for the unbiased identification of anthropogenic contamination within environmental geochemistry, but it cannot provide automated (mathematical) deciphering of the risk element origin.

Distinguishing anthropogenic contamination from coal combustion

Reliable evidence is crucial for differentiating anthropogenic contamination resulting from the coal combustion in the coal basins, as it is demonstrated in this work for the Most Basin. Coal itself can contain higher concentrations of risk elements than local soils; in the Most Basin, this is particularly true for As, Cu, and Pb (Vörös et al. 2019), but the coal dust impact of soils in the Most Basin has never been described. The majority of risk elements, including all those addressed in this paper, resulting from coal combustion can escape from chimneys in the form of fly ash (Ratafia-Brown 1994), unless optimized electrostatic precipitators are used. Lanzerstorfer (2018) and Czech et al. (2020) demonstrated that the finest fly ash particles may contain several hundred mg kg^{-1} of Pb and Zn, as well as tens of mg kg^{-1} of As and Sb, which is approximately an order of magnitude higher than global soil or crustal averages. Stafilov et al. (2018) found in their case study in Macedonia that coal fly ash contains about ten times more Cd, Pb, and Zn than local soils. Analyses of fly ash samples from the Ledvice Power Plants are similar (Supplementary Information, Fig. S8 in Sect. "Sampling and analyses"). Conversely, George et al. (2015) reported that fly ash from a local power plant contained only slightly higher levels of risk elements than local soils.

While dust from coal combustion can indisputably pose risks to living organisms, particularly through inhalation, ingestion, or dermal contact (Zhang et al. 2020a, b; Skála et al. 2022), its contribution to soil composition is limited due to the substantial natural pool of risk elements in soils and the relatively small total mass contribution of fly ash in soils. In other words, emissions in the ploughed soils horizons are substantially "diluted" by natural soil components. Consequently, coal combustion-related contamination is often not evident in risk element soil maps, particularly when the natural variability of soils is adequately considered (Stafilov et al. 2018; da Silva Júnior et al. 2019; Adamec et al. 2024). Estimates of how small can be the fly ash contribution to soil risk elements around the Ledvice Power Plant are presented in Sect. "Sampling and analyses" in Supplementary Information.

The number of studies reporting the harmful impact of coal combustion on soil risk elements continues to grow. However, many of these studies do not take the true complexity of soil chemistry into account. Common shortcomings include: (1) associating coal combustion with soil enrichment by non-volatile elements, whose concentrations in coal are mostly below soil averages (e.g., Cu, Mn, or Ti) (Bhuiyan et al. 2010; Pandey et al. 2016; Feng et al. 2024); (2) utilizing soil data from cities and highly industrialized areas that contain numerous coal-unrelated contamination sources (Hanousková et al. 2021; Rouhani et al. 2023; Feng et al. 2024); and (3) overinterpreting the natural variability of local soils where overall concentrations of risk elements remain below or equal to common background values (Jia et al. 2023; Feng et al. 2024). Moreover, modern technologies such as fly ash capture and desulfurization should help prevent soil contamination by the majority of risk elements.

Contamination from power plants: distance and severity

The majority of risk element emissions from coal combustion occur in fly ash. Coarser particles settle close to their source and can be identified in soil maps. Adamec et al. (2024) found elevated topsoil concentrations of mercury (Hg) and Zn at a distance less than approximately 10 km from the Ledvice Power Plant. Similar ranges (10 to 15 km) have been observed in previous studies, such as in the Hg contamination study by Rodríguez Martín et al. (2016) and analyses of PAHs (Li et al. 2020; Kravchenko et al. 2024). Elevated levels of Cd and Ni within 4 km of a coal-fired power plant were reported in soils in the work of Han et al. (2025). Around the largest Serbian coal-fired power plant "Nikola Tesla A", the highest concentrations of Cd, Co, Fe, Mn, Pb and Zn were at a distance of 1 km, and the concentrations of Ni and Cu gradually increased up to a distance of 4 km (Tanić et al. 2018), however, there were no significant upward increases of risk element contents in soil horizons. For elevated concentrations to serve as evidence of point-source contamination, the contaminated soils should exhibit a spatial distribution around the point source. However, in the Most Basin, spatial analyses have demonstrated this pattern for Cd, Hg, and Zn but not for As and Sb (Adamec et al. 2024).

In majority of documented cases, such expected spatial autocorrelation (contamination in a circle or ellipse around a power plant position) is absent. Instead, contamination was found scattered across the landscape around the power plants (Stafilov et al. 2018; Nanos et al. 2015; Jia et al. 2023; Feng et al. 2024), as if local soil variability was more important than the emission control. In this work, the spatial association of contamination was examined visually in scatterplots in Fig. 6.

The results of this work confirm that elevated concentrations of Cd, PAHs, and Zn within a 10 km radius of the Ledvice Power Plant (indicated by the red line of medians in Fig. 6) can be related to coal combustion. This results for Cd and Zn confirm conclusions by Adamec et al. (2024), who also examined spatial distribution of the risk elements in the Most Basin soils from the official Czech monitoring of the agricultural topsoil contamination. Newly in this work the impact of the coal combustion on the soils was quantified. Table 6 summarizes the estimated impact of the power plant, represented as the difference between the median contaminant concentrations within 10 km from the major power plant in the target area (red line in Fig. 6) and the median in the reference area with similar bedrock but shaded from the powerplant by the mountain ridge (the green line in Fig. 6). The estimated impact is rather modest relative to the local background (Table 6). What is even more important, the current emissions of fly ash from the Ledvice Power Plants, estimated in the Supplementary Information, would not produce the increase in topsoil concentrations of risk elements estimated as the maximal estimate of the coal combustion impact in Table 6, unless the fly as contributed considerable fraction of the topsoils, that is surely not the case in the Most Basin floor. Locally high concentrations of As and Pb in some sites around the Ledvice Power Plant are obviously associated with geogenic anomalies, which also show elevated Sb levels (Fig. 6).

Among the most frequently mentioned contaminants in the Most Basin floor (Vácha et al. 2015), only As exceeds the preventive limit for Czech agricultural soils (Table 6). However, these preventive limits do not represent real risks for biota as some researchers incorrectly supposed (Rouhani et al. 2023); the preventive limits merely prevent farmers from further increasing contaminant concentrations through practices like applying sewage sludge to the soil. Additionally, the highest soil As on the Most Basin floor is of the geogenic origin, as shown above. In the Most Basin, the elevated soil concentrations of As and Pb thus correspond more to the slopes of the Ore Mountains rather than proximity to coal combustion centres (Adamec et al. 2024) and does not represent real risk. This paper also highlights the overlooked impact of local felsic effusive rocks around the Ledvice Power Plant, on which the soils contain high levels of As, Be, and Pb (Tables 1 and 2). Several preceding studies attributed elevated risk element concentrations in the soils of the Most Basin to coal combustion without direct proofs, while some samples were taken from the slopes of the Ore Mountains (Skála et al. 2022) or from cities located just below those slopes (Hanousková et al. 2021). Additionally, this paper points out the possible influence of other historical contamination sources, such as the fuel complex in Litvínov (Fig. 1), where petrol was synthesized from coal during World War II. The synthetic fuel production was targeted during bombardments in that war (Dolejš et al. 2020). The fires of industrial enterprises could contribute significantly to surrounding soil contamination as it is known from the study by Appleton and Cave (2018).

Conclusion

The contribution of coal combustion to soil contamination by risk elements in the Most Basin, Czech Republic, is smaller than it has been believed as still repeated in some current publications. The measurable impacts of coal combustion on soil risk elements from the 770 MW power plant studied here are limited and not toxicologically relevant. This paper highlights critical aspects, particularly the lithogenic origin of soils, which must be considered to understand local natural soil variability—an aspect often overlooked in many previous studies on the impacts of coal combustion on soils. In the Most Basin, the relevant geochemically anomalous rocks have a limited spatial extent; however, their weathering products have spread to surrounding areas during the Pleistocene and Holocene, complicating their accurate interpretation using conventional soil contamination maps. Geogenic anomalies require particular attention—anomalous risk element concentrations cannot be inferred solely from geological maps or calculated from lithogenic element concentrations, which are currently common approaches in soil contamination studies. The unbiased identification of the coal combustion impacts needs careful holistic work respecting natural complexity of the soil chemical variability. The authors express concern that the impacts of coal combustion may have been exaggerated in many previously studied coal basins due to the neglect of the natural factors behind soil composition.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12665-025-12458-z.

Acknowledgements The authors would like to thank P. Kroupová Hailová (UJEP) for her assistance in soil sampling (part of the expanded dataset) and T. Kovářová (UJEP) for her help with PAH extraction and analyses, both as part of their bachelor's projects. The authors acknowledge generous help by R. Beránek (Ledvice Power Plant), who provided the samples of fly ash and information on the fly ash production. P. Vorm and M. Maříková (IIC CAS) conducted sample processing and X-ray fluorescence (XRF) analyses. M. Königová (UJEP) assisted with laboratory analyses for ICP-MS.

Author contributions TMG planned the work and drafted the manuscript. SA performed most of the sampling, worked with AR extracts, conducted granulometry, and finalized the manuscript. ŠT conducted GIS work, while DB contributed to sampling efforts. HB and SK supervised PAH extractions and performed GC-MS analyses. ML conducted PCA and MLRR analyses and edited the manuscript, and IP provided statistical analysis of granulometric curves.

Funding Open access publishing supported by the institutions participating in the CzechELib Transformative Agreement. Analytical work was funded by institutional sources from IIC and UJEP and by the student project UJEP-SGS-2023-44-001-2, granted to S. Adamec. Minor part of analytical works and the last steps in manuscript finishing were funded by RUR project (CZ.10.02.01/00/22_002/0000210).

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

Adamec S, Tůmová Š, Hošek M, Lučić M, Matys Grygar T (2024) Pitfalls of distinguishing anthropogenic and Geogenic reasons for risk elements in soils around coal-fired power plants: from a case study in the Northwestern Czech Republic to general recommendations. J Soils Sediments 24:1274–1288. https://doi.org/10.1007/s11368-024-03726-9

Amorosi A, Guermandi M, Marchi N, Sammartino I (2014) Fingerprinting sedimentary and soil units by their natural metal contents: a new approach to assess metal contamination. Sci Total Environ 500:361–372. https://doi.org/10.1016/j.scitotenv.2014.08.078

Ander EL, Johnson CC, Cave MR, Palumbo-Roe B, Nathanail CP, Lark RM (2013) Methodology for the determination of normal background concentrations of contaminants in english soil. Sci Total Environ 454–455:604–618. https://doi.org/10.1016/j.scitotenv.2013.03.005

Appleton JD, Cave MR (2018) Variation in soil chemistry related to different classes and eras of urbanisation in the London area. Appl Geochem 90:13–24. https://doi.org/10.1016/j.apgeochem.2017.12.024

Armiento G, Barsanti M, Caprioli R, Chiavarini S, Conte F, Crovato C, De Cassan M, Delbono I, Montereali MR, Nardi E, Parrella L, Pezza M, Proposito M, Rimauro J, Schirone A, Spaziani F (2022) Heavy metal background levels and pollution Temporal trend assessment within the marine sediments facing a brownfield area (Gulf of pozzuoli, Southern Italy). Environ Monit Assess 194:814. https://doi.org/10.1007/s10661-022-10480-3

Aytop H, Koca YK, Şenol S (2023) The importance of using soil series–based geochemical background values when calculating the enrichment factor in agricultural areas. Environ Geochem Health 45:6215–6230. https://doi.org/10.1007/s10653-023-01640-6

Bábek O, Matys Grygar T, Faměra M, Hron K, Nováková T, Sedláček J (2015) Geochemical background in polluted river sediments: how to separate the effects of sediment provenance and grain size with statistical rigour? CATENA 135:240–253. https://doi.org/10.1016/j.catena.2015.07.003

Baize D, Sterckeman T Of the necessity of knowledge of the natural pedo-geochemical background content in the evaluation of the contamination of soils by trace elements. The Science of the Total Environment 264,127–139 https://doi.org/10.1016/S0048-9697(00)00615-X

Ballabio C, Panagos P, Lugato E, Huang JH, Orgiazzi A, Jones A, Ugalde OF, Borrelli P, Montanarella L (2018) Copper distribution in European topsoils: an assessment based on LUCAS soil survey. Sci Total Environ 636:282–298. https://doi.org/10.1016/j.scitotenv.2018.04.268

Ballabio C, Jiskra M, Osterwalder S, Borrelli P, Montanarella L, Panagos P (2021) A Spatial assessment of mercury content in the European union topsoil. Sci Total Environ 769:144755. https://doi.org/10.1016/j.scitotenv.2020.144755

Bhuiyan MAH, Parvez L, Islam MA, Dampare SB, Suzuki S (2010) Heavy metal pollution of coal mine-affected agricultural soils in the Northern part of Bangladesh. J Hazard Mater 173:384–392. https://doi.org/10.1016/j.jhazmat.2009.08.085

Breiter K, Durisova J, Hrstka T, Korbelova Z, Hlozkova Vankova M, Vasinova Galiova M, Kanicky V, Rambousek P, Knesl I, Dobes P, Dosbaba M (2017) Assessment of magmatic vs. metasomatic processes in rare-metal granites: a case study of the Cınovec/Zinnwald Sn–W–Li deposit, Central Europe. Lithos 292–293:198–217. https://doi.org/10.1016/j.lithos.2017.08.015

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

Czech T, Marchewicz A, Sobczyk AT, Krupa A, Jaworek A, Śliwiński Ł, Rosiak D (2020) Heavy metals partitioning in fly ashes between various stages of electrostatic precipitator after combustion of different types of coal. Process Saf Environ Prot 133:18–31. https://doi.org/10.1016/j.psep.2019.10.033

da Silva Júnior FMR, Ramires PF, dos Santos M, Seus ER, Soares MCF, Muccillo-Baisch AL, Mirlean N, Baisch PRM (2019) Distribution of potentially harmful elements in soils around a large coal-fired power plant. Environ Geochem Health 41:2131–2143. https://doi.org/10.1007/s10653-019-00267-w

Dolejš M, Samek V, Veselý M, Elznicová J (2020) Detecting world war II bombing relics in markedly transformed landscapes (city of most, Czechia). Appl Geogr 119 https://doi.org/10.1016/j.apgeog.2020102225

Feng F, Zhou Y, Su W, Sun J, Li Y (2024) Homology and heterogeneity of soil trace elements of coal power production bases in arid and semi-arid areas of Northwest China. Environ Monit Assess https://doi.org/10.1007/s10661-024-12738-4

Filzmoser P, Nordhausen K (2020) Robust linear regression for high-dimensional data: an overview. WIREs Comp Stats 13:4:1–18. https://doi.org/10.1002/wics.1524

Filzmoser P, Varmuza K (2009) Introduction to multivariate statistical analysis in chemometrics. CRC

García D, Tapia J, Aguilera J, Vega C, Zúñiga P, Lavin P, Rojas L, Valdés J (2024) Contamination of urban soils in a historical mining town of Northern Chile. Environ Earth Sci. https://doi.org/10.1007/s12665-024-11445-0

George J, Masto RE, Ram LC, Das TB, Rout TK, Mohan M (2015) Human exposure risks for metals in soil near a Coal-Fired Power-Generating plant. Arch Environ Contam Toxicol 68:451–461. https://doi.org/10.1007/s00244-014-0111-x

Greenacre M (2019) Variable selection in compositional data analysis using pairwise logratios. Math Geosci 51:649–682

Han H-J, Song.-W, Yoon D, Lee J-U (2025) Soil pollution with heavy metals in the vicinity of coal-fired power plants in Taean and seocheon, Chungnam province, South Korea. Environ Geochem Health 47(10). https://doi.org/10.1007/s10653-024-02322-7

Hanousková B, Száková J, Rychlíková E, Najmanová J, Košnář Z, Tlustoš P (2021) The risk assessment of inorganic and organic pollutant levels in an urban area affected by intensive industry.

Environ Monit Assess (2021) 193: 68. https://doi.org/10.1007/s10661-020-08825-x

Hošek M, Pavlíková P, Šoltýs M, Tůmová Š, Matys Grygar T (2024) Distinguishing Geogenic load and anthropogenic contribution to soil contamination in mineralised mountain landscape of ore mountains (Czech Republic) using cumulative distribution functions. Land. https://doi.org/10.3390/land13020218

Hubert M, Engelen S (2004) Robust PCA and classification in biosciences. Bioinformatics 20(11):1728–1736. https://doi.org/10.1093/bioinformatics/bth158

Hubert M, Debruyne M, Rousseeuw PJ (2018) Minimum Covariance Determinant and extensions. Wiley Interdisciplinary Reviews: Computational Statistics, 10 (3):e1421. Retrieved from https://wires.onlinelibrary.wiley.com/doi/full/https://doi.org/10.1002/wics.1421

Jia Y, Yang X, Yan X, Duguer W, Hu H, Chen J (2023) Accumulation, potential risk and source identification of toxic metal elements in soil: a case study of a coal–fired power plant in Western China. Environ Geochem Health 45:7389–7404. https://doi.org/10.1007/s10653-023-01661-1

Joliffe IT (1986) Principal component analysis. Springer-, New York

Jung H, Lim D, Xu Z, Jeong K (2016) Secondary grain-size effects on Li and Cs concentrations and appropriate normalization procedures for coastal sediments. Estuar Coast Shelf Sci 175:57–61. https://doi.org/10.1016/j.ecss.2016.03.028

Kravchenko E, Sushkova S, Raza MH, Minkina T, Dudnikova T, Barbashev A, Maksimov A, Wong MH (2024) Ecological and human health impact assessments based on long–term monitoring of soil PAHs near a coal–fired power plant. Environ Geochem Health 46:288. https://doi.org/10.1007/s10653-024-02060-w

Lanzerstorfer C (2018) Fly Ash from coal combustion: dependence of the concentration of various elements on the particle size. Fuel 228:263–271. https://doi.org/10.1016/j.fuel.2018.04.136

Li R, Cheng M, Cui Y, He Q, Guo X, Chen L, Wang X (2020) Distribution of the soil PAHs and health risk influenced by coal usage processes in Taiyuan city, Northern China. Int J Environ Res Public Health 17:6319. https://doi.org/10.3390/ijerph17176319

Lučić M, Mikac N, Vdović N, Bačić N, Dinis P, Milačič R (2023) Distinguishing between natural and anthropogenic sources of potentially toxic elements in sedimentary materials along the Sava river. Appl Geochem 151(2023):105619. https://doi.org/10.1016/j.apgeochem.2023.105619

Matys Grygar T, Popelka J (2016) Revisiting geochemical methods of distinguishing natural concentrations and pollution by risk elements in fluvial sediments. J Geochem Explor 170:39–57. https://doi.org/1016/j.gexplo.2016.08.003

Matys Grygar T, Elznicová J, Kiss T, Smith HG (2016) Using sedimentary archives to reconstruct pollution history and sediment provenance: the Ohře river, Czech Republic. CATENA 144:109–129. https://doi.org/10.1016/j.catena.2016.05.004

Matys Grygar T, Elznicová J, Tůmová Š, Kylich T, Skála J, Hron K, Alvarez-Vázquez MÁ (2023) Moving from geochemical to contamination maps using incomplete chemical information from long-term high-density monitoring of Czech agricultural soils. Environ Earth Sci 82. https://doi.org/10.1007/s12665-022-10692-3

McKinley JM, Hron K, Grunsky EC, Reimann C, de Caritat P, Filzmoser P, van den Boogaart KG, Tolosana-Delgado R (2016) The single component geochemical map: fact or fiction? J Geochem Explor 162:16–28. https://doi.org/10.1016/j.gexplo.2015.12.005

Miller JR (2013) Forensic assessment of metal contaminated rivers in the 21st century using geochemical and isotopic tracers. Minerals 3(2):192–246. https://doi.org/10.3390/min3020192

Murray J, Guzmán S, Tapia J, Nordstrom DK (2023) Silicic volcanic rocks, a main regional source of Geogenic arsenic in waters: insights from the Altiplano-Puna plateau, central Andes. Chem Geol 629:121473. https://doi.org/10.1016/j.chemgeo.2023.121473

Nanos N, Grigoratos T, Rodríguez Martín JA, Samara C (2015) Scale-dependent correlations between soil heavy metals and as around four coal-fired power plants of Northern Greece. Stoch Environ Res Risk Assess 29:1531–1543. https://doi.org/10.1007/s00477-014-0991-3

Okedeyi OO, Nindi MM, Dube S, Awofolu OR (2013) Distribution and potential sources of polycyclic aromatic hydrocarbons in soils around coal-fired power plants in South Africa. Environ Monit Assess 185:2073–2082. https://doi.org/10.1007/s10661-012-2689-7

Otčenášek J, Židov P, Příkaský L (2019) Project of the Academy of Sciences of the Czech Republic: Strategy AV21, program Paměť v digitálním věku 2019 „Přerušená paměť – zaniklá a zanikající podoba českého venkova (Memory in the Digital Age 2019 program Interrupted Memory - the extinct and disappearing form of the Czech countryside). https://prerusenapamet.eu.cas.cz/uhli1.php

Pampura TV, Meili M, Holm K, Candaudap F, Probst A (2019) Buried paleosols as reference objects for assessing the current level of soil pollution with lead in the lower Volga steppes. EURASIAN SOIL Sci 52:34–49. https://doi.org/10.1134/S1064229319010113

Pandey B, Agrawal M, Singh S (2016) Ecological risk assessment of soil contamination by trace elements around coal mining area. J Soils Sediments 16:159–168. https://doi.org/10.1007/s11368-015-1173-8

R Core Team (2023) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. http://www.R-project.org

Rajchl M, Uličný D, Grygar R, Mach K (2009) Evolution of basin architecture in an incipient continental rift: the Cenozoic Most Basin, Eger Graben (Central Europe). Basin Res 21;269–294. https://doi.org/10.1111/j.1365-2117.2008.00393.x

Ramsay J, Silverman BW (2005) Functional data analysis. Springer, New York

Ratafia-Brown JA (1994) Overview of trace element partitioning in flames and furnaces of utility coal-fired boilers. Fuel Process Technol 39:139–157. https://doi.org/10.1016/0378-3820(94)90177-5

Reimann C, de Caritat P, GEMAS Project Team, NGSA Project Team (2012) New soil composition data for Europe and australia: demonstrating comparability, identifying continental-scale processes and learning lessons for global geochemical mapping. Sci Total Environ 416:239–252. https://doi.org/10.1016/j.scitotenv.2011.11.019

Reimann C, Fabian K, Birke M, Filzmoser P, Demetriades A, Négrel P, Oorts K, Matschullat J, de Caritat P, GEMAS Project Team (2018) GEMAS: Establishing geochemical background and threshold for 53 chemical elements in European agricultural soil. Appl Geochem 88:302–318. https://doi.org/10.1016/j.apgeochem.2017.01.021

Rodríguez Martín JA, Nanos N (2016) Soil as an archive of coal-fired power plant mercury deposition. J Hazard Mater 308:131–138. https://doi.org/10.1016/j.jhazmat.2016.01.026

Rouhani A, Gusiatin MZ, Hejcman M (2023) An overview of the impacts of coal mining and processing on soil: assessment, monitoring, and challenges in the Czech Republic. Environ Geochem Health 45:7459–7490. https://doi.org/10.1007/s10653-023-01700-x

Rudnick RL, Gao S (2003) The composition of the continental crust. Treat Geochem 3:1–64. https://doi.org/10.1016/b0-08-043751-6/03016-4

Skála J, Boahen F, Száková J, Vácha R, Tlustoš P (2022) Arsenic and lead in soil: impacts on element mobility and bioaccessibility.

Environ Geochem Health (2022) 44:943–959. https://doi.org/10.1007/s10653-021-01008-8

Skála J, Matys Grygar T, Achasova A (2024) Novel definition of local baseline values for potentially toxic elements in Czech farmland using adaptive spatial weighting. Applied Geochemistry. Volume 170, September 2024, 106082. https://doi.org/10.1016/j.apgeochem.2024.106082

Stafilov T, Šajn R, Arapčeska M, Kungulovski I, Alijagić J (2018) Geochemical properties of topsoil around the coal mine and thermoelectric power plant. J Environ Sci Health Part A. https://doi.org/10.1080/10934529.2018.1445076

Sterckeman T, Douay F, Baize D, Fourrier H, Proix N, Schwartz C, Carignan J (2006) Trace element distributions in soils developed in loess deposits from Northern France. Eur J Soil Sci 57:392–410. https://doi.org/10.1111/j.1365-2389.2005.00750.x

Sun HY, Sun X, Xiao K, Si XB, Zheng MJ, Ye YQ, Hao JH (2024a) In situ analysis of mica from the Daguanshan granites in Western yunnan: implication for Sn enrichment and mineralization. Ore Geol Rev 165:105869. https://doi.org/10.1016/j.oregeorev.2024.105869

Sun Y, Zhao Y, Hao L, Zhao X, Lu J, Shi Y, Ma C, Li Q (2024b) Application of the partial least square regression method in determining the natural background of soil heavy metals: A case study in the Songhua river basin, China. Sci Total Environ 918:170695. https://doi.org/10.1016/j.scitotenv.2024.170695

Tanić MN, Ćujić MR, Gajić BA, Daković MZ, Dragović SD (2018) Content of the potentially harmful elements in soil around the major coal-fired power plant in serbia: relation to soil characteristics, evaluation of Spatial distribution and source apportionment. Environ Earth Sci 77(28). https://doi.org/10.1007/s12665-017-7214-4

Tapia J, Audry S, Townley B, Duprey JL (2012) Geochemical background, baseline and origin of contaminants from sediments in the mining-impacted Altiplano and Eastern cordillera of oruro, Bolivia. GEOCHEMISTRY-EXPLORATION Environ Anal 12:3–20. https://doi.org/10.1144/1467-7873/10-RA-049

Vácha R, Skála J, Čechmánková J, Horváthová V, Hladík J (2015) Toxic elements and persistent organic pollutants derived from industrial emissions in agricultural soils of the Northern Czech Republic. J Soils Sediments 15:1813–1824. https://doi.org/10.1007/s11368-015-1120-8

van den Boogaart KG, Egozcue JJ, Pawlowsky-Glahn V, Australian (2014) New Z J Stat, 54(2):171–194. https://doi.org/10.1111/anzs.12074

Vilà M, Martínez-Lladó X (2015) Approaching Earth surface geochemical variability from representative samples of geological units: the congost river basin case study. J Geochem Explor 148:79–95. https://doi.org/10.1016/j.gexplo.2014.08.013

Vörös D, Geršlová E, Nývlt D, Geršl M, Kuta J (2019) Assessment of Geogenic input into Bilina stream sediments (Czech Republic). Environ Monit Assess 191:114. https://doi.org/10.1007/s10661-019-7255-0

Waroszewski J, Sprafke T, Kabala C, Musztyfaga E, Łabaza B, Woźniczka P (2017) Aeolian silt contribution to soils on mountain slopes (Mt. Ślęża, Southwest Poland). Quatern Res 89:702–717. https://doi.org/10.1017/qua.2017.76

Zhang J, Liu F, Huang H, Wang R, Xu B (2020a) Occurrence, risk and influencing factors of polycyclic aromatic hydrocarbons in surface soils from a large-scale coal mine, huainan, China. Ecotoxicol Environ Saf 192:110269. https://doi.org/10.1016/j.ecoenv.2020.110269

Zhang L, McKinley J, Cooper M, Peng M, Wang Q, Song Y, Cheng H (2020b) A regional soil and river sediment geochemical study in Baoshan area, Yunnan province, Southwest China. J Geochem Explor 217:106557. https://doi.org/10.1016/j.gexplo.2020.106557

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.