

Critical review on source tracing method based on heavy metals in soil: Key parameters, data processing and environmental geochemical information extraction

Qiuhui Yao^a , Yuanhao Wang^b , Guangzhu Zhou^b*, Sai Wu^c and Yongxiang Li^a

^aThe Third Exploration Team, Shandong Bureau of Coal Geology, Tai'an 271000, China

^bCollege of safety and Environmental Engineering, Shandong university of science and technology, Qingdao 266590, China

^cKey Laboratory of Pollution Processes and Environmental Criteria, Ministry of Education, Tianjin Key Laboratory of Environmental Remediation and Pollution Control, College of Environmental Science and Engineering, Nankai University, Tianjin 300071, China

Received: 04/03/2024, Accepted: 30/05/2024, Available online: 21/06/2024

*to whom all correspondence should be addressed: e-mail: zhouguangzhu@sdust.edu.cn <https://doi.org/10.30955/gnj.005871>

Graphical abstract

Abstract

Heavy metals have significant environmental impacts on human health and ecosystems. Appropriate remediation measures can be developed based on the source characteristics of heavy metals. So tracking the source of heavy metal pollution has become a prerequisite for assessing its potential pollution pathways and environmental risks. The stable isotope ratio method is currently widely used for tracing heavy metals. This article focuses on the application of Pb, Hg, Cd, Cu, Zn, and Sr stable isotope techniques in environmental heavy metal tracing. The key parameters, data processing and environmental geochemical information extraction methods of heavy metals in soil are also discussed. Finally, based on the current research status and progress, the source identification of heavy metals in soil is prospected.

Keywords: soil; heavy metal; source apportionment; stable isotope; data mining

1. Introduction

Heavy metal pollution in soil has attracted much attention because of its serious impact on the environment and human health(Rajendran *et al.* 2022). Many public concerns are focusing on the mechanism, prevention, and control technology of soil pollution by heavy metal, in which the source of heavy metal in the soil is the key scientific problem necessary to be resolved before the accurate prevention and control of heavy metal pollution(Hua *et al.* 2018). Heavy metal tracing or apportionment research usually involves work such as basic environmental investigation, sampling, determination, data analysis, and result expression, among which data analysis is an important step in extracting the environmental geochemical information. However, these data analysis methods vary greatly depending on the specific tracing or apportionment technology. In general, there are some common methods for analyzing the sources of soil heavy metal pollutants, e.g., spatial analysis based on geographic information system (GIS) (Guo *et al.* 2018; Zhu *et al.* 2017a), multivariate statistical analysis (Guan *et al.* 2016; Song *et al.* 2018), receptor model (Lang *et al.* 2015; Lu *et al.* 2018), and the stable isotope techniques (Kong *et al.* 2018a; Zhang *et al.* 2018), etc. The stable isotope technology has unique advantages in identifying the relationship between pollutants in acceptors and sources, and then quickly and accurately identifying the source of pollutants (Wang *et al.* 2021a). Therefore, the stable isotope technology is widely used in the tracing or apportionment of heavy metal pollutants in soil and sediments (Figure 1).

Because of the stable composition of Pb isotopes and no fractionation in the secondary process, the Pb stable isotope is frequently adopted in environmental geochemical research. However, with the deepening understanding of the complexity of the environmental pollution sources, the Pb stable isotopes alone can no longer meet the research needs (Yang *et al.* 2019). New stable isotopes, such as Cu, Zn, Cd, Hg, and Sr, were found to be available supplements for better analysis of the type of source members and their contribution ratio in pollution receptor, and reliable methods can be provided

Qiuhui Yao, Yuanhao Wang, Guangzhu Zhou, Sai Wu and Yongxiang Li (2024), Critical review on source tracing method based on heavy metals in soil: Key parameters, data processing and environmental geochemical information extraction, *Global NEST Journal*, **26**(6), 05871.

by them for solving heavy metal pollution apportionment. In recent years, many achievements have been reported in diverse literature. It's necessary to analyze and summary on the new progress in heavy metal tracing based on Pb, Hg, Cu, Zn, Cd, and other isotopes in soil for more widely understanding and application in heavy metal apportionment. In this paper, the application and limitations of different analytical methods of soil heavy metal pollution sources are systematically discussed, mainly focusing on (1) stable isotope tracing (2) Principal component analysis/factor analysis with Multiple Linear Regression, (3) Positive matrix factorization, (4) combination of multiple analytical methods.

2. Overview of Pollution Source Analysis Methods

Contaminant source analysis is a qualitative or quantitative method for studying pollutant sources. This method has been widely used in the source analysis of atmospheric particulate matter, soil and water pollution(Li *et al.* 2020). Pollution source analysis methods can be roughly divided into two categories: one is to qualitatively determine the main source type of pollutants, called source identification. And the other is to quantitatively estimate the contribution rate of each emission source while determining the main pollution source type, which is called source apportionment(Shi *et al.* 2024). It is

usually necessary to combine the two, firstly, the source identification method is used to make a preliminary judgment of the main causes of pollution, and secondly, the source analysis technology method is used to further analyze the source and contribution of pollutants in the soil.

3. Methods for analyzing the sources of heavy metal pollution in soil

There are various models and methods available for analyzing soil heavy metal pollution sources. These methods include isotope ratio analysis, chemical mass balance method, principal component analysis, positive definite matrix factor analysis, and more(Sun *et al.* 2020). However, soil environmental pollution problems are often caused by multiple sources of pollution, which creates difficulties for source apportionment. Single-source apportionment methods may not accurately identify the sources of pollution. Therefore, it is essential to use multiple methods together to improve the accuracy of source apportionment and increase the credibility of source apportionment results(Yang *et al.* 2021). In this paper, the most commonly used methods in the analysis of soil heavy metal sources are summarized, and the similarities and differences of various methods are also compared.

Table1. Index and models of isotopes for heavy metals tracing in soil (Details refer to the table attachment)

3.1. Index and calculation models of stable isotope for heavy metal tracing

Stable isotope measurements can provide "fingerprint" information about the source of metals. Since the stable isotopes were used in environmental geochemistry, the application has evolved from the initial qualitative identification to the quantitative contribution of pollutants sources(Wang *et al.* 2021b). Index (e.g. isotope ratio, fractionation value, etc.) and their combinations are often adopted for calculation and identification (Table 1).

Because the geochemical behavior in the environment is different, the stable isotopes of a certain element, these index and calculation models are also quite different. Table 1 shows some successful cases of this index and models of stable isotopes for Pb, Hg, Cu, Zn, and Cd tracing.

Figure 1. (a) Literature analysis on the identification of soil heavy metal sources (as of September 15, 2023) (b) Symbiotic analysis of heavy metal source identification in soil

Figure 2. Pb Isotope Composition in Soils of Different Regions (redrawn according to (Hu *et al.* 2018a; Kong *et al.* 2018a; Sun *et al.* 2017a; Zhu *et al.* 2017a))

3.1.1. Mechanism and application of Pb isotopes

The relative mass difference of Pb isotopes is trivial, the isotope fractionation effect is weak and usually is ignored, so Pb isotopes are often used as "fingerprint" in pollutants source tracing research (ChengHu 2010). As to Pb, there are four stable isotopes, ²⁰⁴Pb, ²⁰⁶Pb, ²⁰⁷Pb, and ²⁰⁸Pb, and their relative abundances are 1% , 24 $\%$, 23 $\%$ and 52 $\%$ respectively (Shetaya *et al.* 2019), ²⁰⁴Pb is stable and has no radioactive precursor has been found, but the latter three are atomic fission products.

As to the ratio parameters, ²⁰⁶Pb/²⁰⁴Pb, ²⁰⁸Pb/²⁰⁴Pb, $207Pb/204Pb$, $208Pb/206Pb$, and $206Pb/207Pb$ are often used for the identification of the pollution source and the migration process of heavy metal pollutants. Sun *et al.* (2017a) found that the ratio of $^{206}Pb/^{207}Pb$ showed a regular trend in the soil near the traffic roads in Xiamen. The smallest ratio occurred at 0 meters from the highway (1.1839), the farther away from the roadbed the larger the ratio of ²⁰⁶Pb/²⁰⁷Pb was, and it exceeded 1.20 at the distance of 31 meters away from the highway, where was rarely affected by human activities. Hu *et al.* (2018a) found that the ²⁰⁶Pb/²⁰⁷Pb was 1.1645 in the soil near Nanjing, which was similar to the value of ²⁰⁶Pb/²⁰⁷Pb in the dust fall and the fertilizer, obviously disturbed by human activities. The isotopic composition of lead in soil and some sources of some typical cities in China (Figure 2 and Table 2) showed the same law described above. There are certain differences in the ratio of ²⁰⁸Pb/²⁰⁶Pb in soil from different cities in China, mainly due to the different characteristics of land plots and the specific reasons for pollution, e.g., the soil in Nanjing was mainly affected by dust fall and fertilizers, while the soil in Guilin was affected by tailings and geological background, the soil in Beijing Iron and Steel Plant was affected by production dust, coal, and geological background, but the soil in Xiamen was affected by automobile emission.

Combined with Pb stable isotopes and geostatistical methods, the contribution percentage of natural Pb and anthropogenic Pb in soil can be calculated, and the changes of Pb pollutants in environmental media can be well understood, so as to take accurate and effective measures to prevent its pollution(Wei *et al.* 2021). In recent years, more and more multi-source data analysis methods have been used to study the source of heavy metals in soil, such as Hu *et al.* (2018a) used a comprehensive analysis method combining positive matrix factor analysis, input flux analysis, spatial analysis, and Pb isotope analysis to analyze the sources of heavy metals in agricultural land around Nanjing, and the pollution information could be supported by each other, and more accurately description of the main sources of heavy metals in soil was provided. Huang *et al.* (2015) used a combination of spatial analysis, clustering, principal component analysis, and Pb stable isotopes to analyze the sources of heavy metals in agricultural land around Wenling City, Zhejiang Province, and found that the Pb pollution in the soil was due to automobile sources and solid waste, and there was also a small amount of fertilizer pollution. Zhu *et al.* (2017a) used a variety of statistical analysis methods combined with stable isotope analysis to analyze the sources of heavy metals in the soil of abandoned industrial areas and its surrounding area and used a ternary hybrid model of the Pb stable isotope to analyze the relative contribution of dust generated by steel production, coal combustion, natural background and other pollution end elements, which provide a good guide for the quantitative analysis of pollution sources. Based on a multi-source isotope mixing model of lead isotopes in lichen and road sediments, Dietrich *et al.* (2021) studied the sources of lead pollution in complex urban environments and discovered the dominant position of industrial pollution in typical medium-sized cities in the United States and the legacy issues of leaded gasoline. The commonly used binary or ternary mixed models for stable isotopes in existing research can only calculate the contribution rates of two or three pollutant sources, and the contribution of more sources can be analyzed based on the Isosource model(Huang *et al.* 2015) or Bayesian hybrid model (Kong *et al.* 2018a). By comparing the Pb isotope ratios of the road sweeping waste with different pollution sources, Nawrot *et al.* (2020) reported that Pb enrichment came from the combustion of coal, excluding Pb sources for gasoline and diesel. Huang *et al.* (2020) used the MC-ICP-MS method to determine lead isotope ratios from 45 soil samples and potential sources of contamination. Through analysis, it was found that the pollution level of Pb and Cd in the soil on both sides of the Yellow River was mild-to-moderate,

Table 2. Pb isotopic composition of pollution terminals in different regions (Details refer to the table attachment)

At present, the traceability research on Pb stable isotope technology has attracted much attention. Due to the low variability of lead isotopes in natural and industrial processes, it is possible to trace pollution sources using lead isotopes. In recent years, researchers have used the lead isotope ratio method to visually indicate the source of pollutants and calculate the contributions of different lead pollution sources. This method has made a significant contribution to indicating the source of pollutants.

3.1.2. Mechanism and application of Hg isotope traceability

There are 7 stable isotopes of Hg $(^{196}$ Hg, 198 Hg, 199 Hg, 200 Hg, 201 Hg, 202 Hg, and 204 Hg), and they have a relative mass difference of 4%. There are isotope fractionation effects of Hg in some physical and chemical processes (such as redox chemical processes, and volatilization

processes of 0-valent Hg). The measurable changes existed in the isotope composition of the acceptors and pollution contributors in different processes. Hg can be used as an effective traceability tool to identify the fingerprint of Hg sources in the environment, and to detect important reactions affecting Hg migration, conversion, toxicity, and bioaccumulation in natural systems(FoucherHintelmann 2006).

There are isotope fractionation effects of Hg in some physical and chemical processes (such as redox chemical processes, and volatilization processes of 0-valent Hg). The measurable changes existed in the isotope composition of the acceptors and pollution contributors in different processes. Hg stable isotopes show both massdependent fractionation (MDF) and mass-independent fractionation (MIF)(BlumBergquist 2007). Among them, MDF is common in environmental physical, chemical, and

biological processes, while MIF may be triggered by the Magnetic isotope effect (MIE), photochemical process, and nuclear volume effect (NVE).

The typical composition of some Hg stable isotopes was collected from the literature and shown in Figure 3. From the stable isotope composition of Hg in different sites of the world, HintelmannLu (2003) found that $δ²⁰²Hg$ varies from -1.33‰ to 0‰ in different Hg ore cinnabar samples around the world. BlumBergquist (2007) studied the Hg isotopic composition of cinnabar samples from the world's largest Hg mine and found that δ^{202} Hg had an average value of -0.54‰. Other researchers found that the average value of $δ²⁰²Hg$ in the Wanshan Hg ore cinnabar sample in Guizhou Province was 0.74‰ 0.11‰(Yin *et al.* 2013). The above is Hg isotope MDF. However, the presence of Hg isotopes in coal is not related to fractionation. Feng *et al.* (2013) analyzed previously reported data on Hg isotopes in coal and found that δ^{202} Hg was between -0.11‰ and -2.98‰, with a mean value of --1.29‰±0.64‰, which was lower than the isotopic value in the cinnabar sample in Hg ore. The article also pointed out that the average value of MIF of Hg isotopes with Δ^{199} Hg is -0.09 ‰ \pm 0.17 ‰.

Figure 3. Hg Isotope Composition in Different Receptor (Drawn according to literature (Biswas *et al.* 2008; Feng *et al.* 2010; Feng *et al.* 2013; Smith *et al.* 2008; Yin *et al.* 2013; Zhang *et al.* 2018))

In recent years, the research on mercury pollution in the environment has attracted widespread attention. For example, Feng *et al.* (2010) used Hg stable isotope technology to find that the isotopic composition of Hg in the sediment of the Hongfeng Reservoir was similar to the isotopic composition of Hg in eroded soil and atmospheric dust, so as to judge that the source of Hg pollution in the sediment was soil and atmospheric dust precipitation. Zhang *et al.* (2018) used spatial analysis and Hg stable isotope technology to analyze the pollution sources of Hg in industrial zones in Jiangsu Province, and found that the value of $Δ^{202}$ Hg in contaminated soil was similar to the value of $Δ^{202}$ Hg in atmospheric dust (Figure 3). The above two results show that atmospheric dust precipitation is one of the sources of Hg, indicating that the analysis results of Hg pollution sources are reliable. Therefore, Hg stable isotope technology has a good application prospect in the traceability of pollutants.

The isotope ratio of Hg can be used to trace mercury pollution sources, which has a good effect on controlling mercury pollution. At the same time, allocation can also be made based on the contribution of MDF and MIF of Hg to pollution in nature.

3.1.3. Mechanism and application of Cd isotope traceability

There are 8 stable isotopes of Cd in nature, ¹⁰⁶Cd, ¹⁰⁸Cd, $110Cd$, $111Cd$, $112Cd$, $113Cd$, $114Cd$, $116Cd$, and their abundances are respectively $1.25\%, 0.89\%, 12.8\%, 12.8\%$,24.1%,12.2%,28.7%and 7.49%(Zhong *et al.* 2020). It has long been reported that the stable isotope composition of Cd was reported, but the testing technology at that time was not advanced enough, and the measured isotopic composition changes were on the same order of magnitude as the deviation of the analytical measurement method, so the true changes in isotopes and their mechanisms could not be found. With the emergence and maturity of MC-ICP-MS analysis methods, the measurement and application of transition metal Cd isotopes are possible (Peng *et al.* 2021). Cd isotopes are new geochemical tracers, and studying the fractionation of Cd isotopes is an important step in their traceability applications. In this paper, soil samples were collected as well as the Cd isotopic composition in the contaminated terminal element (Figure 4). Among them, the δ^{114} Cd/¹¹⁰Cd value in the polluted soil in the mining area is small, and the δ^{114} Cd/¹¹⁰Cd value in the agricultural soil is relatively heavy, and the soil in different study areas is affected by different pollutants, so different δ^{114} Cd/ 110 Cd values will be produced.

Figure 4. Cd Isotope Composition in Soil and Polluted endmember in Different Areas (Quoted from literature (Cloquet *et al.* 2006; Gao *et al.* 2008; Salmanzadeh *et al.* 2017; Yang *et al.* 2019; Zhang *et al.* 2013)).

Cd is relatively toxic and belongs to the transition metal element. The use of Cd stable isotope technology can analyze the source of Cd pollution so that effective measures can be quickly taken to solve the pollution

problem. Salmanzadeh *et al.* (2017) studied 66 years of field experiments in New Zealand and found that fertilizer from 2,000 years ago was the main source of Cd, based on this observation, they concluded that there has not been a large amount of Cd in the soil in recent years. It was found that the isotopic assignment of Cd sources in soil was caused by the application of phosphorus source rocks and fertilizers in 1998, and the influence of end elements containing different Cd isotopes on Cd sources in soil was determined by the Bayesian mixed model. In addition, by analyzing the fractionation of the isotope Cd in industrial processes, Cloquet *et al.* (2006) can track the contamination of Cd, determine the isotope and concentration of Cd in the surface soil, and then determine that the Cd in the soil may come from industrial dust, slag, and agriculture. The source and contribution of Cd are expressed through the relationship between the value in the surface soil and the Cd concentration, and the higher the Cd concentration, the greater the contribution. Yang *et al.* (2019) used Cd stable isotopes and other geochemical tools to study the differences between Cd stable isotopes in rivers and sediments affected by acidic mineral water, as well as changes in Cd stable isotopes along a section of the river. As the river moves downstream, the value of δ^{114} Cd^{/110}Cd in the section near the middle and lower reaches of farmland suddenly increases, and the heavy isotopic value is similar to the isotopic value in agricultural fertilizer and water samples (Figure 4), indicating that agricultural fertilizer and agricultural wastewater may be the second source of Cd pollution in the water body, thus indicating that Cd stable isotope technology can not only analyze the source of pollutants in environmental media but also trace the migration and transformation pathways of Cd in the environment. Wen *et al.* (2015) used Cd stable isotopes and Pb stable isotopes to analyze the sources of heavy metal pollution in the soil of the Pingding mining area in Yunnan, China, and found that the average δ^{114} Cd/ 110 Cd value in the background soil in this area was 0.41‰. The average ²⁰⁶Pb/²⁰⁷Pb value is 1.1902. However, the values in the contaminated soil samples differed, with δ ¹¹⁴Cd/¹¹⁰Cd values varying from -0.59‰ to 0.33‰ and ²⁰⁶Pb/²⁰⁷Pb values ranging from 1.1764 to 1.1896. Finally, according to the isotopic relationship between different pollution end elements and soil samples without sampling points, the different pollution sources in the southwest and north-east directions were better divided. Liao *et al.* (2022) used the alteration index and exchange state information of Cd in soil to study the weathering process of Cd in soil. At the same time, the isotopic composition of soil Cd was determined to determine the geogenesis isotope trend of soil Cd in the weathering process and the influence of human input on soil surface, and it was found that there was great variability in the mass transfer process of Cd in soil, indicating the importance of artificial input Cd, and anthropogenic source was an important factor causing serious cadmium pollution in the urbanrural junction of Chengdu. Yan *et al.* (2021) analyzed the Cd content and Cd isotopic composition in rice plants (roots, stems, leaves and grains) and topsoil, as well as

possible sources of pollution (agricultural fertilizers, industrial dusts and automobile exhaust). The results showed that all three pollution sources contributed to the Cd content of rice grain, with industrial dust contributing the most to Cd pollution.

Through the isotope ratio method, cadmium from human and natural sources can be distinguished, and Cd pollution from human sources can be traced. The use of this method can effectively reduce the harm caused by cadmium pollution.

3.1.4. Mechanism and application of Zn, Cu, Sr isotope traceability

Zn, Cu and Sr are essential trace elements for life activities, but they become toxic elements at high concentrations, so it is of great significance to understand their environmental geochemical behavior in detail, and their stable isotopes provide new ideas for studying their geochemical behavior in the environment. Their stable isotopes in nature are ⁶³Cu, ⁶⁵Cu, ⁸⁴Sr, ⁸⁶Sr, ⁸⁸Sr, ⁶⁴Zn, ⁶⁶Zn, ⁶⁷Zn, ⁶⁸Zn, ⁷⁰Zn, and their abundances in nature are 69.17%, 30.83%, 0.56%, 9.86%, 6.98%, 82.95%, 48.63%, 27.90%, 4.10%, 18.75%, and 0.62%. Zn isotope fractionation is influenced by the thermodynamic and kinetic effects of the source region composition, and is produced by mineral fractionation, re-deposition and leaching, and precipitation. Therefore, the application of Zn isotopes can trace the Zn of different reaction processes. At present, the composition of the Zn isotopes obtained varies greatly, the δ^{66} Zn value of terrestrial materials varies from -0.91‰ to 1.7‰, and the δ^{66} Zn of Zn-bearing ores of different types of deposits in the world varies between -0.43‰ and 0.47‰(Zhu *et al.* 2013) (Figure 5). Cu isotope fractionation is influenced by biological, redox and precipitation processes. The composition of Cu isotopes varies between different receptors, so the source of Cu contamination can be traced by using Cu stable isotopes(Viers *et al.* 2023). The concentration of Sr in the environment is low, widely distributed, and the ${}^{87}Sr/{}^{86}Sr$ ratio is mainly affected by the source of Sr, and the value of $87Sr/86Sr$ in different

receptors is different, so Sr isotopes can be used to trace the source of Sr pollution.

In recent years, many scholars have used Zn, Cu and Sr stable isotopes to trace anthropogenic and natural sources of heavy metal pollution in soil. Bigalke *et al.* (2010) used Cu and Zn stable isotope techniques to study the sources of Cu and Zn in soil near the Slovak copper smelter, changes in different soil depths, and geochemical behavior, and found that the soil did not change much with δ^{65} Cu in most wastes, so source identification could not be carried out. With the increase of soil depth to 0.4m, δ^{65} Cu becomes lighter, which may be caused by the equilibrium reaction between dissolved and adsorbed Cu species during soil transport by melting Cu. At the same time, through the stable isotope change of Zn, it was analyzed that Zn may have undergone biogeochemical fractionation of soil and plant systems and equilibrium reaction fractionation between dissolved Zn and adsorbed Zn in soil, and due to the evaporation kinetic fractionation of Zn, the isotopic composition of Zn of natural substances is significantly different from that of Zn produced by smelting, so that the source of Zn pollution in soil can be judged. This is consistent with the results of Fekiacova *et al.* (2015)'s analysis of light isotope signatures of volatile pollutants emitted from plants and heavy isotope signatures in solid waste, so Zn stable isotopes are potentially valuable in identifying the source of contaminants in soil. Juillot *et al.* (2011) used Zn stable isotopes to analyze soil samples near a Pb-Zn smelter in northern France and found that δ^{66} Zn in deep soil had a mean value of 0.31 \pm 0.38‰, representing the background value of Zn, while the δ^{66} Zn value in the soil surface layer ranged from 0.38 \pm 0.45‰ to 0.76 \pm 0.14‰, representing Zn of anthropogenic origin (Figure 5). Matthies *et al.* (2014) analyzed the Zn isotopes in the waste slag leachate and found that the Zn isotope ratio does not change with concentration, so the Zn isotope can be used to trace the Zn source from the mine. DesaultyPetelet-Giraud (2020) discusses the effects of isotope fractionation that can alter the initial characteristics of various Zn sources, and discusses the effects of various anthropogenic Zn emissions based on the study of zinc isotopes.

The study of a heavy metal isotope can no longer meet the needs of problem solving, and we need to use a variety of isotopes to trace the source of heavy metal pollution to achieve satisfactory results. Sun *et al.* (2017a) used Pb stable isotopes and Sr stable isotopes to resolve the sources of contaminants in soil. Among them, the pollution sources of Pb and Sr in the topsoil sample drift from automobile exhaust emissions or coal combustion to the parent material as the sampling point is away from the road. For soil profiles, with the increase of soil depth, the source of Pb and Sr in the soil shifts from human activities to natural sources. This result is consistent with the analytical results obtained using Pb and Sr stable isotope techniques. Kong *et al.* (2018a)'s analysis of the slightly contaminated soil profile in the study area also showed that as the soil depth increased, the source of pollution was the migration of human activities to natural sources. The combination of Pb and Sr stable isotopes can better trace heavy metal pollution in soil. Huang *et al.* (2015) used Hg and Pb to stabilize isotopes to trace heavy metal contaminants in soil with good results. Therefore, stable isotope combination technology is an important development direction for tracing heavy metal pollution, and the use of multiple stable isotopes to comprehensively analyze the sources of heavy metal pollution will better provide reliable decisions for the control and prevention of heavy metal pollution.

Combining the isotope ratio method of multiple elements can more accurately trace pollution sources. Due to the different pollution sources of non copper heavy metals, combining multiple heavy metal isotopes for analysis can accurately indicate the source of pollutants.

3.2. Principal component analysis/factor analysis with Multiple Linear Regression (PCA/FA-MLR)

Principal component analysis/factor analysis with Multiple Linear Regression is a method that uses principal component analysis or factor analysis in combination with multiple linear regression to quantify the source of pollutants. That is, the acceptor data is first dimensionally reduced through principal component analysis or factor analysis, the relationship between multiple variables is analyzed, fewer representative factors, that is, pollution factors, are extracted, and then these factors are subjected to multiple linear regression to the elements of the receptor to obtain the regression coefficient of the multiple linear regression equation, which can reflect the contribution value of these factors to the receptor(Fraino 2023). The multiple linear regression model can be represented by the following equation.

$$
y = \sum_{i=1}^{p} m_i X_i + b \tag{Eq.1}
$$

y is the dependent variable, *p* is the number of principal factors (pollution sources) extracted by PCA/FA, and *m*ⁱ is the standard regression coefficient. *X*ⁱ is the *i*th principal factor score *b* is the regression constant.

The researchers used principal component analysis to analyze the sources of chromium, nickel, copper, zinc and lead contamination in Dehui City (Sun *et al.* 2013). The rotational component matrix shows that Zn, Ni, and Cr are related to the first component (F1). The main element of the second component (F2) is Pb, and the third component (F3) includes Cu. Since the variability of Cr, Ni and Zn appears to be controlled by the parent material, F1 can be considered as a diagenetic component witch can explains 43% of the total variance; and F2 can explain 25% of the total, and common sources of lead are automobile exhaust, industrial exhaust, so F2 can be considered a source of industrial pollution; F3 contains Cu, which can be defined as a man-made component associated with agricultural practices. In order to improve the accuracy of tracing pollution sources, the likely location of pollution sources can be analyzed by combining PCA with other analytical models. For example, in a multifunctional industrial park in Anhui Province, Zeng *et al.* (2022) used a

combination of principal component analysis (PCA), geographic detector (Geodetector) and multiple linear regression distance (MLRD). Principal component analysis (PCA) and geographic detector are used to determine the possible location and type of the source, and the source assignment of different elements is done by distance multiple linear regression (MLRD). A comparative analysis of the regression equations of Model 1 and Model 3 derived from principal component analysis shows that rivers, roads and mines are the main sources of pollution.

3.3. Positive matrix factorization (PMF)

Positive matrix factorization (PMF) is a factor analysis method based on the least squares method, which decomposes the matrix without negative constraints and can be optimized by using the standard deviation of the data. The basic principle of the positive matrix factorization 2D model is to split the pollutant content matrix into two matrices, namely the source component matrix and the source contribution matrix. The residual matrix consists of the difference between the pollutant content matrix (actual data) and the split source component matrix and source contribution matrix (analyzed data).The PMF model factorizes the original matrix $X(i \times j)$, decomposing it into two factor matrices, $F(k\times i)$ and $G(i\times k)$, and a residual matrix $E(i\times i)$, as shown below:

$$
X_{ij} = \sum_{k=1}^{p} G_{ik} F_{kj} + E_{ij}
$$
 Eq.2

where X_{ij} represents the concentration of element i in sample j; F_{kJ} is the concentration of the j element in source k; p represents the number of sources, that is, the source component spectral matrix; *Gik*'s contribution to the i-sample, the contribution rate matrix of the source; *E*ij is the source and k is the residual matrix.

Figure 6. PMF model using soil heavy metal data (Liang *et al.* 2017) (a) Factor profiles from PMF model using soils heavy metals concentrations data. (b) Correlation coefficients between different heavy metals. (c) Factor contributions of heavy metals calculated by PMF model.

Before using Positive Matrix Decomposition (PMF) to identify source of heavy metals, it is necessary to distinguish the pollution sources. Therefore, studying the correlation between different heavy metals and the main sources of heavy metals is requisite. A review of the literature found that high levels of As, Pb and Hg may be related to coal combustion through atmospheric deposition and industrial production (Bhuiyan *et al.* 2015). Studies have shown that the main sources of Hg in China's soil are agricultural activities, coal mining and, non-ferrous metal smelting accounting for15%, 39% and 46% respectively, while the metallurgical industry and municipal solid waste are also the main sources of Cd (Liang *et al.* 2015). The researchers used positive definite matrix factor decomposition (PMF) to analyze the spatial distribution characteristics and sources of heavy metals in surface soil in Lianyuan City, a typical coal mining city(Liang *et al.* 2017). As shown in Figure 6, they calculated the contribution rate of each heavy metal source based on the factor fingerprint of each heavy metal. The sources of heavy metals in the soil of Lianyuan City were natural sources, atmospheric sedimentation, industrial activities and agricultural activities, and the contribution distribution was 33.61%, 26.05%, 23.44% and 16.91%. In general, the influence of human factors on the metal element content of Lianyuan soil was dominant, accounting for about 66.39%(Wu *et al.* 2020). In recent years, researchers have analyzed the heavy metal content of agricultural soils in Tianjin, China based on positive matrix decomposition (PC-PMF) based on partition calculation. Through the analysis, it is concluded that sludge application, atmospheric deposition and irrigation are the main anthropogenic sources, and the contribution rates are 26.60%, 19.56% and 2.86%, respectively. In this study, the source of heavy metals in soil in the suburban agricultural area of Kaifeng City was quantitatively analyzed by combining Pb-Zn-Cu isotopic composition, PMF model and UNTEX (Chen *et al.* 2022).

3.4. A variety of analysis methods combined to analyze pollution sources

As the problem of environmental pollution becomes more and more complex, the single-source analysis method will not meet the requirements of accurate quantitative analysis of pollution sources, and when carrying out soil source analysis research in a specific area, it can be considered to select an appropriate source analysis method based on the actual local situation, or combine a variety of methods to improve the accuracy and availability of source analysis results. Due to the different assumptions or principles of different soil source analysis methods, showing their respective limitations and advantages in application, the obtained source analysis results may be different, and finding out the reasons for the differences is helpful to comprehensively and objectively interpret the source analysis results.

To identify sources of contamination in several subdivisions of the mine, the researchers established a precise technique for environmental forensics using multiple isotopes (Cu, Zn, Pb, and S) and PMF. At the same time, this technique can also be used to understand fractionation and potential geochemical processes in the mine that affect the ratio of Cu, Zn and S isotopes. Using this method, it was found that the δ^{65} Cu value of soil

samples downstream of tunnel seepage was significantly higher than that in other areas (Kim *et al.* 2023). Due to the large number of isotope fractionation during oxidative dissolution, the ore sulfide leads to a significant increase in δ^{65} Cu in the tunnel permeate (Wall *et al.* 2011). Therefore, δ^{65} Cu is particularly useful for distinguishing sulfide (tailings) and water (flat seepage) contamination. At the same time, the source of contaminants was identified by PMF analysis and the contamination of As and Cu-enriched tailings was distinguished. The impact of downstream copper-rich tailings showed that the proportion of chalcopyrite was lower than that of toxic sand, which affected the distribution of soil pollution. This method of tracing PMF and stable isotopes has also been used to study the hazards caused by mining to agricultural systems and human health around mining sites(Liu *et al.* 2022). In addition, PMF, PCA and spatial analysis can be combined to evaluate the sources of heavy metals, calculate the contribution rate of anthropogenic pollution sources to heavy metal pollution, and speculate on the possible anthropogenic pollution sources (Sun *et al.* 2019).

4. Summary

Due to the spatial heterogeneity of environmental media, there are differences in the composition of heavy metal stable isotopes in different polluted areas, and there are differences in the stable isotopes of heavy metals of different terminal element pollutants in the same pollution area, at present, the composition spectrum of heavy metal stable isotopes in various regions has not been fully established, we need to further improve the heavy metal isotope composition of pollutant sources, in order to determine the source of pollutants in environmental media more quickly and accurately, and to regularly determine the isotopes This enables rapid response to pollution incidents and the ability to remediate and monitor their source.

The problem of pollution in environmental media is becoming more and more complicated, and the use of only one heavy metal pollutant traceability method is likely to bias the analysis results, or make them meaningless, and ultimately fail to meet our expected requirements. When conducting source analysis research on areas contaminated by heavy metals, we should consider the geographical, hydrological, meteorological and other conditions of the research area, and compare the analysis results with the help of geostatistical analysis methods, multivariate statistical analysis, stable isotope techniques, etc., to find out the differences, so as to improve the accuracy and recognition of the analysis results.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Qiuhui Yao, Yuanhao Wang*,* GuangzhuZhou, Sai Wu and Yongxiang Li. The first draft of the manuscript was written by Qiuhui Yao, Yuanhao Wang and all authors commented on previous versions of the

manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Special Fund for Scientific Research of Shandong Bureau of Coal Geology (Grant numbers MTDZKY-2024-25).

Availability of data and materials

The authors confirm that the data supporting the findings of this study are available within the article.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose

References

- Bhuiyan MAH, Dampare SB, Islam MA, Suzuki S (2015) Source apportionment and pollution evaluation of heavy metals in water and sediments of buriganga river, bangladesh, using multivariate analysis and pollution evaluation indices. Environ Monit Assess 187:21.
- Bigalke M, Weyer S, Kobza J, Wilcke W (2010) Stable cu and zn isotope ratios as tracers of sources and transport of cu and zn in contaminated soil. Geochim Cosmochim Ac 74:6801-6813.
- Biswas A, Blum JD, Bergquist BA, Keeler GJ, Xie Z (2008) Natural mercury isotope variation in coal deposits and organic soils. Environ Sci Technol 42:8303-8309.
- Blum JD, Bergquist BA (2007) Reporting of variations in the natural isotopic composition of mercury. Anal Bioanal Chem 388:353-359.
- Chen Z, Ding Y, Jiang X et al (2022) Combination of unmix, pmf model and pb-zn-cu isotopic compositions for quantitative source apportionment of heavy metals in suburban agricultural soils. Ecotox Environ Safe 234:113369.
- Cheng H, Hu Y (2010) Lead (pb) isotopic fingerprinting and its applications in lead pollution studies in china: a review. Environ Pollut 158:1134-1146.
- Cloquet C, Carignan J, Libourel G, Sterckeman T, Perdrix E (2006) Tracing source pollution in soils using cadmium and lead isotopes. Environ Sci Technol 40:2525-2530.
- Desaulty A, Petelet-Giraud E (2020) Zinc isotope composition as a tool for tracing sources and fate of metal contaminants in rivers. Sci Total Environ 728:138599.
- Dietrich M, Krekeler MPS, Kousehlar M, Widom E (2021) Quantification of pb pollution sources in complex urban environments through a multi-source isotope mixing model based on pb isotopes in lichens and road sediment. Environ Pollut 288:117815.
- Fekiacova Z, Cornu S, Pichat S (2015) Tracing contamination sources in soils with cu and zn isotopic ratios. Sci Total Environ 517:96-105.
- Feng X, Foucher D, Hintelmann H, Yan H, He T, Qiu G (2010) Tracing mercury contamination sources in sediments using

mercury isotope compositions. Environ Sci Technol 44:3363-3368.

- Feng X, Yin R, Yu B, Du B (2013) Mercury isotope variations in surface soils in different contaminated areas in guizhou province, china. Chinese Science Bulletin 58:249-255.
- Foucher D, Hintelmann H (2006) High-precision measurement of mercury isotope ratios in sediments using cold-vapor generation multi-collector inductively coupled plasma mass spectrometry. Anal Bioanal Chem 384:1470-1478.
- Fraino PE (2023) Using principal component analysis to explore multi-variable relationships. NATURE REVIEWS EARTH & ENVIRONMENT 4:294.
- Gao B, Liu Y, Sun K et al (2008) Precise determination of cadmium and lead isotopic compositions in river sediments. Anal Chim Acta 612:114-120.
- Guan Q, Wang L, Wang F et al (2016) Phosphorus in the catchment of high sediment load river: a case of the yellow river, china. Sci Total Environ 572:660-670.
- Guo A, Cui J, Xu X, Ma X (2018) Spatial distribution of soil nutrients based on gis and geostatistics. Chinese Agricultural Science Bulletin 34:72-79.
- Hintelmann H, Lu S (2003) High precision isotope ratio measurements of mercury isotopes in cinnabar ores using multi-collector inductively coupled plasma mass spectrometry. The Analyst 128:635.
- Hu W, Wang H, Dong L et al (2018a) Source identification of heavy metals in peri-urban agricultural soils of southeast china: an integrated approach. Environ Pollut 237:650-661.
- Hu W, Wang H, Dong L et al (2018b) Source identification of heavy metals in peri-urban agricultural soils of southeast china: an integrated approach. Environ Pollut 237:650-661.
- Hua C, Zhou G, Yin X et al (2018) Assessment of heavy metal in coal gangue: distribution, leaching characteristic and potential ecological risk. Environ Sci Pollut R 25:32321-32331.
- Huang Y, Li T, Wu C et al (2015) An integrated approach to assess heavy metal source apportionment in pen-urban agricultural soils. J Hazard Mater 299:540-549.
- Huang Y, Zhang S, Chen Y et al (2020) Tracing pb and possible correlated cd contamination in soils by using lead isotopic compositions. J Hazard Mater 385:121528.
- Juillot F, Maréchal C, Morin G et al (2011) Contrasting isotopic signatures between anthropogenic and geogenic zn and evidence for post-depositional fractionation processes in smelter-impacted soils from northern france. Geochim Cosmochim Ac 75:2295-2308.
- Kim D, Kwon H, Im D (2023) Determination of contamination sources and geochemical behaviors of metals in soil of a mine area using cu, pb, zn, and s isotopes and positive matrix factorization. J Hazard Mater 447:130827.
- Kong J, Guo Q, Wei R et al (2018a) Contamination of heavy metals and isotopic tracing of pb in surface and profile soils in a polluted farmland from a typical karst area in southern china. Sci Total Environ 637:1035-1045.
- Kong J, Guo Q, Wei R et al (2018b) Contamination of heavy metals and isotopic tracing of pb in surface and profile soils in a polluted farmland from a typical karst area in southern china. Sci Total Environ 637:1035-1045.
- Lang Y, Li G, Wang X, Peng P (2015) Combination of unmix and pmf receptor model to apportion the potential sources and

contributions of pahs in wetland soils from jiaozhou bay, china. Mar Pollut Bull 90:129-134.

- Lee P, Kang M, Jeong Y, Kwon YK, Yu S (2020) Lead isotopes combined with geochemical and mineralogical analyses for source identification of arsenic in agricultural soils surrounding a zinc smelter. J Hazard Mater 382:121044.
- Li Y, Yuan Y, Sun C et al (2020) Heavy metals in soil of an urban industrial zone in a metropolis: risk assessment and source apportionment. Stoch Env Res Risk a 34:435-446.
- Liang J, Feng C, Zeng G et al (2017) Spatial distribution and source identification of heavy metals in surface soils in a typical coal mine city, lianyuan, china. Environ Pollut 225:681-690.
- Liang J, Hua S, Zeng G et al (2015) Application of weight method based on canonical correspondence analysis for assessment of anatidae habitat suitability: a case study in east dongting lake, middle china. Ecol Eng 77:119-126.
- Liao R, Ratié G, Shi Z et al (2022) Cadmium isotope systematics for source apportionment in an urban-rural region. Appl Geochem 137:105196.
- Liu L, Xu X, Han J et al (2022) Heavy metal(loid)s in agricultural soils in the world's largest barium-mining area: pollution characteristics, source apportionment, and health risks using pmf model and cd isotopes. Process Saf Environ 166:669-681.
- Lu X, Hu W, Huang B et al (2018) Source apportionment of heavy metals in farmland soils around mining area based on unmix model. environmental science 39:1421-1429.
- Matthies R, Sinclair SA, Blowes DW (2014) The zinc stable isotope signature of waste rock drainage in the canadian permafrost region. Appl Geochem 48:53-57.
- Mattielli N, Petit JCJ, Deboudt K et al (2009) Zn isotope study of atmospheric emissions and dry depositions within a 5 km radius of a pb-zn refinery. Atmos Environ 43:1265-1272.
- Nawrot N, Wojciechowska E, Rezania S, Walkusz-Miotk J, Pazdro K (2020) The effects of urban vehicle traffic on heavy metal contamination in road sweeping waste and bottom sediments of retention tanks. Sci Total Environ 749:141511.
- Peng H, He D, Guo R et al (2021) High precision cadmium isotope analysis of geological reference materials by double spike mc-icp-ms. J Anal Atom Spectrom 36:390-398.
- Rajendran S, Priya TAK, Khoo KS et al (2022) A critical review on various remediation approaches for heavy metal contaminants removal from contaminated soils. Chemosphere 287:132369.
- Salmanzadeh M, Hartland A, Stirling CH et al (2017) Isotope tracing of long-term cadmium fluxes in an agricultural soil. Environ Sci Technol 51:7369-7377.
- Shetaya WH, Marzouk ER, Mohamed EF, Bailey EH, Young SD (2019) Chemical and isotopic fractionation of lead in the surface soils of egypt. Appl Geochem 106:7-16.
- Shi B, Meng J, Wang T, Li Q, Zhang Q, Su G (2024) The main strategies for soil pollution apportionment: a review of the numerical methods. Journal of environmental sciences (China) 136.
- Sivry Y, Riotte J, Sonke JE et al (2008) Zn isotopes as tracers of anthropogenic pollution from zn-ore smelters the riou mort - lot river system. Chem Geol 255:295-304.
- Smith CN, Kesler SE, Blum JD, Rytuba JJ (2008) Isotope geochemistry of mercury in source rocks, mineral deposits

and spring deposits of the california coast ranges. Earth Planet Sc Lett 269:339-407.

- Song Y, Li H, Li J et al (2018) Multivariate linear regression model for source apportionment and health risk assessment of heavy metals from different environmental media. Ecotox Environ Safe 165:555-563.
- Sun C, Liu J, Wang Y, Sun L, Yu H (2013) Multivariate and geostatistical analyses of the spatial distribution and sources of heavy metals in agricultural soil in dehui, northeast china. Chemosphere 92:517-523.
- Sun J, Hu G, Yu R, Lin C, Wang X, Huang Y (2017a) Human health risk assessment and source analysis of metals in soils along the g324 roadside, china, by pb and sr isotopic tracing. Geoderma 305:293-304.
- Sun J, Hu G, Yu R, Lin C, Wang X, Huang Y (2017b) Human health risk assessment and source analysis of metals in soils along the g324 roadside, china, by pb and sr isotopic tracing. Geoderma 305:293-304.
- Sun L, Guo D, Liu K et al (2019) Levels, sources, and spatial distribution of heavy metals in soils from a typical coal industrial city of tangshan, china. Catena 175:101-109.
- Sun X, Wang H, Guo Z et al (2020) Positive matrix factorization on source apportionment for typical pollutants in different environmental media: a review. Environ Sci-Proc Imp 22:239-255.
- Viers J, Freydier R, Grande JA et al (2023) The use of copper isotopes for understanding metal transfer mechanisms within the continuum mine-river-dam (huelva region, spain). Environ Sci Pollut R 30:53275-53294.
- Wall AJ, Mathur R, Post JE, Heaney PJ (2011) Cu isotope fractionation during bornite dissolution: an in situ x-ray diffraction analysis. Ore Geol Rev 42:62-70.
- Wang L, Jin Y, Weiss DJ et al (2021a) Possible application of stable isotope compositions for the identification of metal sources in soil. J Hazard Mater 407:124812.
- Wang L, Jin Y, Weiss DJ et al (2021b) Possible application of stable isotope compositions for the identification of metal sources in soil. J Hazard Mater 407:124812.
- Wei W, Li X, Ling S, Wu X, Liao X (2021) Heavy metal(loid) and pb isotope compositions of black shale weathering profiles on the northern yangtze platform: insights into geochemical behavior, contamination assessment, and source apportionment. Environ Sci Pollut R 28:50230-50244.
- Wen H, Zhang Y, Cloquet C, Zhu C, Fan H, Luo C (2015) Tracing sources of pollution in soils from the jinding pb – zn mining district in china using cadmium and lead isotopes. Appl Geochem 52:147-154.
- Wu J, Li J, Teng Y, Chen H, Wang Y (2020) A partition computingbased positive matrix factorization (pc-pmf) approach for the source apportionment of agricultural soil heavy metal contents and associated health risks. J Hazard Mater 388:121766.
- Yan Y, Sun Q, Yang J, Zhang X, Guo B (2021) Source attributions of cadmium contamination in rice grains by cadmium isotope composition analysis: a field study. Ecotox Environ Safe 210:111865.
- Yang W, Ding K, Zhang P et al (2019) Cadmium stable isotope variation in a mountain area impacted by acid mine drainage. Sci Total Environ 646:696-703.
- Yang X, Yang Y, Wan Y, Wu R, Feng D, Li K (2021) Source identification and comprehensive apportionment of the

accumulation of soil heavy metals by integrating pollution landscapes, pathways, and receptors. Sci Total Environ 786.

- Yin R, Feng X, Wang J et al (2013) Mercury speciation and mercury isotope fractionation during ore roasting process and their implication to source identification of downstream sediment in the wanshan mercury mining area, sw china. Chem Geol 336:72-79.
- Zeng W, Wan X, Wang L, Lei M, Chen T, Gu G (2022) Apportionment and location of heavy metal(loid)s pollution sources for soil and dust using the combination of principal component analysis, geodetector, and multiple linear regression of distance. J Hazard Mater 438:129468.
- Zhang Y, Wang M, Huang B, Akhtar MS, Hu W, Xie E (2018) Soil mercury accumulation, spatial distribution and its source identification in an industrial area of the yangtze delta, china. Ecotox Environ Safe 163:230-237.
- Zhang Y, Zhou Q, Zhu C et al (2013) Cadmium lsotopic fractionation during the supergenic weathering and leaching process and its implication. Earth and Environment 41:612-617.
- Zhao Z, Zhang W, Li X et al (2015) Atmospheric lead in urban guiyang, southwest china: isotopic source signatures. Atmos Environ 115:163-169.
- Zhong Q, Zhou Y, Tsang DCW et al (2020) Cadmium isotopes as tracers in environmental studies: a review. Sci Total Environ 736.
- Zhu G, Guo Q, Xiao H, Chen T, Yang J (2017a) Multivariate statistical and lead isotopic analyses approach to identify heavy metal sources in topsoil from the industrial zone of beijing capital iron and steel factory. Environ Sci Pollut R 24:14877-14888.
- Zhu G, Guo Q, Xiao H, Chen T, Yang J (2017b) Multivariate statistical and lead isotopic analyses approach to identify heavy metal sources in topsoil from the industrial zone of beijing capital iron and steel factory. Environ Sci Pollut R 24:14877-14888.
- Zhu X, Wang Y, Yan B et al (2013) Developments of nontraditional stable isotope geochemistry. Bulletin of Mineralogy, Petrology and Geochemistry 32:651-688.