

Soil Heavy Metal Distribution Characteristics and Source Analysis Model Selection and Soil Pollution Source Types

Yicheng Song^{1,a,*}

¹*Dalian University of Technology, Panjin City, China*

a. 2903587548@qq.com

**corresponding author*

Abstract: Rabindranath Tagore described the world's unfairness to him with the words, "The world has kissed my soul with its pain, asking for its return in songs." Despite the adversity, he responded with kindness. Similarly, the current state of soil environment can be likened to this sentiment. The importance of soil to us is self-evident, yet a multitude of soil environmental problems arise one after another. Soil heavy metal source analysis is the most efficient means to judge and mitigate soil pollution problems. Although the development of soil pollution source analysis models is quite mature, the types of soil pollution source analysis are very diverse, and the regions involved are uneven. In this study, the main heavy metal pollution corresponding to different soil pollution sources and the analysis model of the dominant soil pollution sources corresponding to different pollution sources were reviewed, and the optimal pollution source analysis method should be selected to deal with different soil heavy metal pollution problems. This study aims to better understand the relationship between soil heavy metal distribution characteristics, the selection of source analysis models, and the types of soil pollution sources, so that more efficient use of better soil pollution source analysis models and targeted selection of soil pollution source analysis models for detecting specific soil heavy metals can be achieved.

Keywords: Soil pollution source analysis, Soil heavy metal distribution characteristics, Types of soil pollution sources

1. Introduction

In China, it is estimated that about 150,000 tons of agricultural products are contaminated by Cd each year, including 50,000 tons of rice (accounting for 0.025% of global rice production) [1]. In 2014, the Ministry of Environmental Protection of China, in conjunction with the Ministry of Land and Resources, released the "National Soil Pollution Survey Bulletin" [2], which shows that the condition of our soil environment is not optimistic. The rate of land use exceeding standards in heavily polluted enterprises reached 36.1%, and the rate of cadmium (Cd) exceeding standards was 7.0% [3]. Identifying pollution sources and quantifying their contribution to soil contamination has always been a key issue in the field of soil pollution source analysis. Commonly used receptor model software developed for this purpose includes methods such as Principal Component Analysis (PCA), Positive Matrix Factorization (PMF) model, and Chemical Mass Balance, among others. Receptor models are often diverse. In the work of soil pollution source analysis, most researchers employ the Positive Matrix Factorization model (PMF) and methods based on the combination of PMF with

other techniques Some researchers also use the method of (PCA) and improved methods based on PCA, such as the APCS-MLR model. The diversity of technical methods can lead to difficulties in selecting the appropriate approach for specific problems. To determine the optimal solution for the selection of soil pollution source analysis methods, this study combines examples of using soil source analysis models to analyze soil pollution sources and the corresponding detected soil heavy metal pollutants. The objectives of this study are: (1) to determine the different major heavy metal pollutants corresponding to different pollution sources by reviewing the relationship between the distribution characteristics of soil heavy metals and specific typical pollution sources; (2) to summarize the optimal solutions for selecting soil pollution source analysis models for different soil pollution sources by reviewing the advantages and disadvantages of the models chosen in existing studies, in conjunction with specific soil pollution sources.

2. Literature Review

The PMF (Positive Matrix Factorization) model is suitable for the analysis of complex data sources with high resolution and no prior information [4]. PCA (Principal Component Analysis) is used for dimensionality reduction and data structure simplification, helping to identify principal components. APCS-MLR (Atmospheric Pollution Characterization Study—Multiple Linear Regression) combines source contributions with regression analysis, providing strong quantitative analysis capability, and is suitable for studies that require precise quantification of pollution sources [5]. The UNMIX model is effective in analyzing complex, multi-source mixed data, and is widely applied in atmospheric pollution source apportionment [6].

2.1. The distribution characteristics of heavy metals in soil correspond to different pollution sources.

The increasing severity of soil heavy metal pollution has posed greater potential risks and harm to society. In order to better understand the distribution characteristics of soil heavy metal pollution, it is essential to first recognize its corresponding relationship with different pollution sources. This review summarizes the distribution characteristics of soil heavy metals associated with various pollution sources, as shown in Table 1:

Table 1: The distribution characteristics of soil heavy metals associated with various pollution sources

The main source of pollution	Distribution characteristics of heavy metals in soil	Distribution of major soil heavy metal pollutants	The number of documents and their sources
Industrial sources	Cu, Zn, Ni, Pb, Cd, Hg, Cr	1.Pb, Zn, Cu	6 [8][19][21] [23][24][27]
		2.Cu, Zn, Ni, Pb, Cd, Hg, Cr	
		3.Mn, Co, Ni, Zn, As, Pb	
		4.Cd, Pb	
		5.Hg, Cd	
		6.Hg, As, Cr, Ni	

Table 1: (continued).

Anthropogenic sources	Pb, Zn, Cu, As, Cd, Cr, Hg	1.Pb, Zn, Cu	7 [8][9][10] [22][23][25] [26]
		2.Cd, As	
		3.As, Cd, Pb, Cu	
		4.Hg	
		5.Cd, Pb	
		6.Cd, Zn	
		7.Hg	
Agricultural sources	Pb, Zn, Cu, As, Cd, Cr, Ni, Mn, Co, Hg	1.Cd, As	8 [9][12][15] [17][20][21] [24][27]
		2.Cd, Zn	
		3.As, Cd, Cr	
		4.Pb	
		5.Cd	
		6.Cd, As, Mn, Co, Ni, Zn	
		7.Hg, Cd	
Natural source	As, Hg, Cd, Pb, Zn, Cu, Ni, Cr, Co, Mn	8.Cu, Zn, Ni, Cd, Pb	8 [10][11][13] [15][19][20] [21][26]
		1.As, Cd, Pb, Cu	
		2.Cd, As, Zn, Pb	
		3.Cr, Ni	
		4.Cr	
		5.Cu, Zn, Ni, Pb, Cd, Hg, Cr,	
		6.Cd	
Mixed sources	Cd, As, Zn, Pb, Cr, Hg	7.Mn, Co, Ni, Zn, As, Pb	4 [11][13] [22][24]
		8.As, Co, Cr, Cu, Mn, Ni, Zn	
		1.Cd, As, Zn, Pb	
		2.Hg	
Traffic sources	As, Cr, Cd, Hg, Pb	3.Cd, Pb, Zn	4 [12][13] [24][27]
		4.Hg, Cd	
		1.Pb, Zn	
		2.Cd, Pb	
Mining sources	As, Cr, Cd, Hg, Pb, Zn, Ni, Mn	3.Hg, Cd	2 [13][28]
		4.Pb, Zn, As	
		1.Cr	
		2.Cr, Zn, Ni, Mn	

2.2. Comparison of the efficiency of the analytical model of soil pollution sources under different pollution sources

Soil heavy metal pollution source analysis models have received extensive attention in recent years. Scholars from different fields have developed various source analysis methods through different ways of thinking. However, each source analysis method has its own area of advantage, and most of the literature aims to compare the selection of source analysis models in a single region and the utilization rate of different source analysis models. Few people have focused on the relationship between different pollution sources and the efficiency of the chosen pollution source analysis models. Here, we discuss the efficiency issues of soil heavy metal pollution source analysis models based on some studies.

The UNMIX and PMF models showed good consistency and were also outstanding in terms of source contribution, with high accuracy. The isotope tracer technique is often used in conjunction with related source analysis methods rather than as a standalone source analysis technology. Similar to this case, the comparison of the effectiveness of other pollution source analysis models and the types of pollution sources in their respective regions is shown in the Table 2.(In the table, ">" indicates that the methods on both sides of the symbol were used together in the literature but the method on the left is superior to the one on the right, "=" indicates that the methods on both sides of the symbol were used together in the literature and each has its own advantages and disadvantages, "<" indicates that the methods on both sides of the symbol were used together in the literature but the method on the right is superior to the one on the left.)

Table 2: The comparison of the effectiveness of other pollution source analysis models and the types of pollution sources in their respective regions

The main source of pollution	Source resolution model used	Comparison of the efficiency of the source resolution model	The number of documents and their sources
Industrial sources	UNMIX PMF Isotope labeling APCS-MLR	1.UNMIX>PMF	9 [8][14][16] [18][19][21] [23][24][27]
		2.PMF>APCS-MLR	
		3.APCS-MLR	
		4.PMF=PCA=APCS-MLR=UNMIX	
		5.PMF	
		6.PMF(SOM)	
		7.PMF(SOM)	
		8.PMF	
		9.PMF	
Anthropogenic sources	UNMIX PMF Isotope labeling PCA APCS-MLR	1.UNMIX>PMF	8 [8][9][10] [14][22][23] [25][26]
		2.PCA	
		3.PMF>PCA	
		4.PMF>APCS-MLR	
		5.PMF=PCA	
		6.PMF	
		7.PC-PMF>PMF	
		8.EPAPMF、WALSPMF	
Agricultural sources	PCA PMF Multivariate statistical analysis APCS-MLR CMB, UNMIX	1.PCA	9 [9][12][15] [16][17][20] [21][24][27]
		2.PMF>Multivariate statistical analysis	
		3.PMF=APCS-MLR>CMB	
		4.APCS-MLR	
		5.PMF=UNMIX>PCA	
		6.PMF	
		7.PMF(SOM)	
		8.PMF(SOM)	
		9.PMF	

Table 2: (continued).

Natural source	PMF PCA CA Multivariate statistical analysis APCS-MLR CMB SOM	1.PMF>PCA	
		2.PMF>PCA=CA	
		3.PMF>Multivariate statistical analysis	10
		4.PMF=PCA	[10][11][12]
		5.PMF>APCS-MLR	[13][14][15]
		6.PMF=APCS-MLR>CMB	[19][20][21]
		7.PMF	[26]
		8.PMF	
		9.PMF(SOM)	
		10.EPAPMF、WALSPMF	
Mixed sources	PMF PCA CA APCS-MLR, UNMIX	1.PMF>PCA=CA	4
		2.PMF=PCA	[11][13][18]
		3.PMF=PCA=APCS-MLR=UNMIX	[22]
		4.PMF=PCA	
Traffic sources	PMF、Multivariate statistical analysis、 APCS-MLR	1.PMF>Multivariate statistical analysis	4
		2.APCS-MLR	[12][16][24]
		3.PMF(SOM)	[27]
		4.PMF	
Mining sources	PMF, PCA	1.PMF=PCA	2
		2.PMF	[13][28]
Transportation sources	PMF、PCA、SOM	1.PMF=PCA	2
		2.PMF(SOM)	[13][24]

3. Results

3.1. The soil heavy metal distribution characteristics corresponding to different pollution sources

To objectively compare the different soil heavy metal pollutants primarily contained in different pollution sources, Table 1 lists in detail the main pollutants analyzed by different researchers. We can categorize and rank the primary pollutants. Cadmium (Cd) and Mercury (Hg), as major soil heavy metal pollutants, are mostly found in anthropogenic sources, mixed sources, and agricultural sources. This includes the misuse of pesticides and fertilizers in agricultural activities, as well as the irrigation of farmland with wastewater. The heavy metal pollutant Chromium (Cr), as a primary pollutant, is frequently associated with mining sources. Elements such as Arsenic (As), Zinc (Zn), and Copper (Cu) appear in most pollution sources but are not the primary pollutants. For example, from anthropogenic and natural sources, the main soil heavy metal pollutants are more varied due to the wide range of sources.

3.1.1. Selection of PMF corresponding to different pollution sources.

According to different types of pollution sources, most researchers have chosen to use only the Positive Matrix Factorization (PMF) method for source apportionment and have made improvements to the original PMF method. For example, they have adopted methods such as segmented calculation-PMF and EPA-PMF and have combined the PMF method with technologies like

Self-Organizing Maps (SOM), Geographic Information Systems (GIS), etc. This approach is also the most widely used method for soil pollution source analysis across various types of pollution sources.

3.1.2. Selection of other source apportionment models corresponding to different pollution sources.

In addition to the currently most widely used and efficient PMF for analyzing soil pollution sources, there are also a variety of methods including, UNMIX, PCA, APCS-MLR, and multivariate statistical analysis. Among these, the application of UNMIX and PCA in soil pollution source detection related to industrial sources is also acceptable and can even be superior to PMF to some extent. In the detection of soil pollution sources related to agricultural sources, methods such as PCA and UNMIX are also selected, but they are not superior to the PMF method. In agricultural, traffic, and natural sources, methods like multivariate statistical analysis and isotope labeling have been employed, but their detection efficiency and accuracy are not as good as the aforementioned methods, particularly PCA and PMF. It is worth noting that a significant number of researchers prefer to use multiple methods simultaneously for detection, which can compensate for the shortcomings of different methods. For example, conducting tests using a combination of PMF, PCA, and UNMIX methods.

4. Discussion

According to the different soil pollution source analysis methods corresponding to different pollution sources, the PMF method consistently demonstrates its high efficiency and accuracy regardless of the type of pollution source. Compared to other pollution source analysis methods, PMF is more specific, not only identifying more pollution sources but also distinguishing between different land use types and heavy metal-related sources under different production conditions [13]. Additionally, methods derived from PMF, such as EPAPMF, are more targeted, especially in natural sources. Furthermore, in anthropogenic and industrial sources, the UNMIX model is also widely accepted by researchers for its relative simplicity and ease of use. In natural and mixed sources, models like PCA and APCS-MLR are commonly used. However, both CA and PCA methods are unable to quantitatively analyze the contribution of pollution sources, and the APCS-MLR model has limitations in identifying sources and explaining complex pollution issues. In most cases, these methods need to be used in conjunction with other detection methods or improved to achieve the desired objectives. For example, Grouped Principal Component Analysis (GPCA) is an improved PCA method that combines the advantages of factor analysis and principal component analysis. Having a thorough understanding of the main soil heavy metal elements corresponding to specific pollution sources can make the detection more targeted. Recognizing the specific pollution sources corresponding to each model allows for targeted selection based on the potential pollution in the study area, which can efficiently conduct soil pollution source analysis. However, this study has its limitations; since most researchers tend to use the PMF method, the PMF soil pollution source analysis method appears in the research review regardless of the type of pollution source. This may lead to conclusions that have a certain subjectivity related to the researchers. Nevertheless, it is undeniable that PMF is indeed a relatively efficient and targeted method for analyzing soil heavy metal pollution sources at present.

5. Conclusions

This study aims to review how to distinguish between soil heavy metal contaminant elements and how to select efficient soil pollution source analysis methods in relation to soil pollution sources. Cd and Hg are mostly found in anthropogenic, mixed, and agricultural sources, while Cr is commonly detected in mining sources. Elements such as As, Zn, and Cu are present in most pollution sources but are not the primary contaminants. PMF is currently the most popular and efficient method for soil

pollution source analysis. The application of UNMIX and PCA in industrial sources is also acceptable and may even be superior to PMF to some extent. In the analysis of soil pollution sources related to agricultural sources, methods like PCA and UNMIX are also chosen, but they are not superior to the traditional PMF method. This study is of great significance for the selection of soil pollution source analysis methods in the future, as more scholars can make targeted method selections based on this research for potential pollution source types that may arise.

References

- [1] Wang M Y, Chen A K, Wong M H, et al. Cadmium accumulation in and tolerance of rice (*Oryza sativa* L.) varieties with different rates of radial oxygen loss[J]. *Environmental Pollution*, 2011, 159(6): 1730-1736.
- [2] Chen Nengchang, Zheng Yujie, He Xiaofeng, et al. Analysis of the "Bulletin on the National Survey of Soil Pollution Status".[J]. *Journal of Agro-Environment Science*, 2017, 36(9)..
- [3] Ministry of Environmental Protection, Ministry of Land and Resources. Bulletin on the National Survey of Soil Pollution Status[J]. *China Environmental Protection Industry*. 2014, 36(5): 10 – 11.
- [4] Guan Q, Wang F, Xu C, et al. Source apportionment of heavy metals in agricultural soil based on PMF: A case study in Hexi Corridor, northwest China[J]. *Chemosphere*, 2018, 193: 189-197.
- [5] Greenacre M, Groenen P J F, Hastie T, et al. Principal component analysis[J]. *Nature Reviews Methods Primers*, 2022, 2(1): 100.
- [6] Li Y, Kuang H, Hu C, et al. Source apportionment of heavy metal pollution in agricultural soils around the Poyang Lake region using UNMIX model[J]. *Sustainability*, 2021, 13(9): 5272.
- [7] Bracher C, Frossard E, Bigalke M, et al. Tracing the fate of phosphorus fertilizer derived cadmium in soil-fertilizer-wheat systems using enriched stable isotope labeling[J]. *Environmental pollution*, 2021, 287: 117314.
- [8] Chen Z, Ding Y, Jiang X, et al. Combination of UNMIX, PMF model and Pb-Zn-Cu isotopic compositions for quantitative source apportionment of heavy metals in suburban agricultural soils[J]. *Ecotoxicology and Environmental Safety*, 2022, 234: 113369.
- [9] Zhang S, Liu Y, Yang Y, et al. Trace elements in soils of a typical industrial district in Ningxia, northwest China: Pollution, source, and risk evaluation[J]. *Sustainability*, 2020, 12(5): 1868.
- [10] Anaman R, Peng C, Jiang Z, et al. Identifying sources and transport routes of heavy metals in soil with different land uses around a smelting site by GIS based PCA and PMF[J]. *Science of the Total Environment*, 2022, 823: 153759.
- [11] Wang Y, Guo G, Zhang D, et al. An integrated method for source apportionment of heavy metal (loid) s in agricultural soils and model uncertainty analysis[J]. *Environmental Pollution*, 2021, 276: 116666.
- [12] Liu H, Anwar S, Fang L, et al. Source apportionment of agricultural soil heavy metals based on PMF model and multivariate statistical analysis[J]. *Environmental Forensics*, 2024, 25(1-2): 40-48.
- [13] Liu Z, Wang L, Yan M, et al. Source apportionment of soil heavy metals based on multivariate statistical analysis and the PMF model: A case study of the Nanyang Basin, China[J]. *Environmental Technology & Innovation*, 2024, 33: 103537.
- [14] Zhang M, Wang X, Liu C, et al. Quantitative source identification and apportionment of heavy metals under two different land use types: Comparison of two receptor models APCS-MLR and PMF[J]. *Environmental Science and Pollution Research*, 2020, 27: 42996-43010.
- [15] Jianfei C, Chunfang L, Lixia Z, et al. Source apportionment of potentially toxic elements in soils using APCS/MLR, PMF and geostatistics in a typical industrial and mining city in Eastern China[J]. *Plos one*, 2020, 15(9): e0238513.
- [16] Li N, Li Y, Wang G, et al. The sources risk assessment combined with APCS/MLR model for potentially toxic elements in farmland of a first-tier city, China[J]. *Environmental Science and Pollution Research*, 2022, 29(33): 50717-50726.
- [17] Guan Q, Zhao R, Pan N, et al. Source apportionment of heavy metals in farmland soil of Wuwei, China: Comparison of three receptor models[J]. *Journal of Cleaner Production*, 2019, 237: 117792.
- [18] Jain S, Sharma S K, Mandal T K, et al. Source apportionment of PM10 in Delhi, India using PCA/APCS, UNMIX and PMF[J]. *Particuology*, 2018, 37: 107-118.
- [19] Kuerban M, Maihemuti B, Waili Y, et al. Ecological risk assessment and source identification of heavy metal pollution in vegetable bases of Urumqi, China, using the positive matrix factorization (PMF) method[J]. *PloS one*, 2020, 15(4): e0230191.
- [20] Liu L, Xu X, Han J, et al. 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[J]. *Process Safety and Environmental Protection*, 2022, 166: 669-681.

- [21] Bhuiyan M A H, Karmaker S C, Bodrud-Doza M, et al. Enrichment, sources and ecological risk mapping of heavy metals in agricultural soils of dhaka district employing SOM, PMF and GIS methods[J]. *Chemosphere*, 2021, 263: 128339.
- [22] Song Q Q, Xu X B, Wu Q Y, et al. Quantitative analysis of environmental risk of heavy metal sources in soil based on PMF model[J]. 2022.
- [23] Xia F, Zhao Z, Niu X, et al. Integrated pollution analysis, pollution area identification and source apportionment of heavy metal contamination in agricultural soil[J]. *Journal of Hazardous Materials*, 2024, 465: 133215.
- [24] Xu Y, Wang X, Cui G, et al. Source apportionment and ecological and health risk mapping of soil heavy metals based on PMF, SOM, and GIS methods in Hulan River Watershed, Northeastern China[J]. *Environmental Monitoring and Assessment*, 2022, 194(3): 181.
- [25] Wu J, Li J, Teng Y, et al. A partition computing-based positive matrix factorization (PC-PMF) approach for the source apportionment of agricultural soil heavy metal contents and associated health risks[J]. *Journal of Hazardous Materials*, 2020, 388: 121766.
- [26] Lv J, Wang Y. PMF receptor models and sequential Gaussian simulation to determine the quantitative sources and hazardous areas of potentially toxic elements in soils[J]. *Geoderma*, 2019, 353: 347-358.
- [27] Li X, Liu H, Meng W, et al. Accumulation and source apportionment of heavy metal (loid) s in agricultural soils based on GIS, SOM and PMF: A case study in superposition areas of geochemical anomalies and zinc smelting, Southwest China[J]. *Process Safety and Environmental Protection*, 2022, 159: 964-977.
- [28] Singh S, Maiti S K, Raj D. An approach to quantify heavy metals and their source apportionment in coal mine soil: a study through PMF model[J]. *Environmental Monitoring and Assessment*, 2023, 195(2): 306.