Analysis of Pollution Sources of Heavy Metal in Farmland Soils Based on Positive Matrix Factorization Model

Sheng Yuan¹, Sheng Zhang¹, Yingcheng Sun³, Miao Guo²-³*

¹Chongqing Academy of Eco-Environmental Sciences, Chongqing 401147, China
²Nanjing Soil Research Institute, Chinese Academy of Sciences, Nanjing 210008, China
³Sichuan Geological Engineering Survey Institute Group Co., Ltd, Chengdu 610036, China

*Corresponding author e-mail: miao_guo03@cas.cn

Abstract. In this research, a total of 103 samples of surface farmland soil in Jiangjin district are collected by 1000m × 1000m grid method, and the content distribution characteristics of such heavy metals as Cd, Hg, Zn, Pb, Cu, as, Ni, Cr in the soil are analyzed. Positive Matrix Factorization (PMF) model is adopted to analyze the pollution sources and contribution rates of these eight heavy metals in the study. The results show that the contribution rate of atmospheric deposition source is 19.7%, that of industrial source is 27.0%, that of natural parent material source is 35.5% and that of agricultural source is 17.8%. It can be seen from the analytical results of PMF model that human factors (64.5%) such as industrial and agricultural production is the main pollution source of heavy metal in farmland soil.

Keywords: Farmland Soil, Heavy Metal, PMF Model, Source Analysis

1. Introduction
In the field of environmental protection, "heavy metal" refers to the general term of metals and metalloids (such as arsenic) with greater toxicity (Duffus 2002, Tchounwou, Yedjou et al. 2012). Because they have the characteristics of concealment, hysteresis, stability and difficult degradation to soil pollution, they are considered as important environmental pollutants (Nicholson, Smith et al. 2003).

Receptor models for contaminated areas, such as Chemical Mass Balance (CMB), Principal Component Analysis/Absolute Principal Components (PCA/APCs), Positive Matrix Factorization (PMF) model and Unmix are effective methods for quantitative source contribution, but they are mainly used in quantitative source analysis of air pollution (Li, Jang et al. 2003, Lee, Liu et al. 2008, Liu, Chen et al. 2009). PMF model simplifies high-dimensional variables by using correlation matrix and covariance matrix, and transforms them into several comprehensive factors. Such method does not need to measure complex original spectrum, it can not only limit the elements and share rate of decomposition matrix, but also deal with the missing and imprecise data. It is a new source solution technology with simple and effective operation (Chen and Lu 2018, Guan, Wang et al. 2018). In this
paper, PMF model is used to analyze heavy metals in the typical farmland soil of Jiangjin District, the old industrial base of Chongqing, so as to provide support for the control and treatment of heavy metals pollution.

2. Materials and Methods

2.1 Introduction to the Study Area
Jiangjin district lies in the middle and upper reaches of the Yangtze River, and is at the end of the Three Gorges Reservoir area. It is located between east longitude of 105° 49’ - 106° 38’ and north latitude of 28° 28’ - 29° 28’. Jiangjin district is surrounded by mountains and hills. The landform is mainly composed of hills with low mountains, which are divided into terraces, hills and mountains, with hills accounting for 78.2% and others accounting for 21.8%. And the soil type is mainly composed of purple soil mainly developed in new alluvium (Qh), old alluvium (Qp) and in Suining Formation (J2sn).

As an important industrial base in Chongqing, Jiangjin district has four major industrial parks and more than 1200 enterprises.

2.2 Sample Collection and Experimental Analysis
1000m × 1000m grid method is used to sample the surface soil of the suburban farmland from 0 to 20 cm in the study area. A total of 103 surface soil samples are collected. The soil samples in each grid are collected using the flower-shaped "grid method", and 5 sub-samples are collected, and each sub sample point collects about 200g of surface soil. The five sub sample points are mixed evenly as the analysis sample of the sample point, and the sample weighs about 1 kg, all sample points are precisely located by GPS.

After the soil sample is air-dried, it is sampled by the quartet method, mixed with a 20-mesh sieve, and then spread on a plastic film. About 20g of the soil sample is randomly obtained and ground by agate mortar, so that it passes through 100 mesh sieve, and then put into the self sealed bag to prepare for the graphite furnace of nitric acid perchlorate hydrofluoric acid system to digest the soil sample. The content of heavy metals such as Cd, Pb, Cu, Zn, Ni, Cr is determined by inductively coupled plasma mass spectrometry (ICP-MS), and that of As, Hg is determined by inductively coupled plasma atomic emission spectrometry (CPAES). Three parallel samples are set for each sample. The national standard soil samples GSS-4 and GSS-6 are used for the statistics and analysis of the soil heavy metal content data. Excel2013 and SPSS19.0 are used for the source analysis, and EPA PMF50 principal component analysis and PSS19.0 software principal component analysis is adopted to calculate score coefficient of the factor.

2.3 PMF Model
2.3.1 Principle of PMF model
The PMF model is a source analysis method based on factor analysis technology. It was first proposed by Platero et al. in 1994. In this study, PMF5.0 is used to analyze the source of heavy metals. The model decomposes the original matrix X into two factor matrices gk and f and the residual matrix e. The formula is shown in the following.

\[
x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij}
\]  

(1)

Where, \(X_{ij}\) represents the concentration of the j-th heavy metal in sample i; \(g_{ik}\) is the contribution rate of the k-th pollution source in sample i; \(f_{kj}\) is the characteristic value of the pollution source k to the j-th heavy metal concentration. And the residual matrix \(e_{ij}\) is calculated by the minimum value of the self-standard function Q.
\[ Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( \frac{e_{ij}}{u_{ij}} \right)^2 \]  

(2)

In the formula, \( u_{ij} \) refers to the uncertainty of the \( j \)-th heavy metal in the sample \( i \).

### 2.3.2 Calculation of uncertainty

The selection of uncertainty calculation algorithms is mainly based on commonly used calculation methods. This study selects two commonly used uncertainty algorithms in the analysis of soil heavy metal sources, which can be seen in Table 1.

The above two uncertainty calculation methods can be summarized as: \( U_{nc} = \sqrt{a^2 + b^2}, \text{Unc}=a+b \). The correlation between parameter \( a \) and data standard deviation or particle concentration is called multiplication term, while the correlation between parameter \( b \) and method detection limit is called additive term. Both algorithm 1 and algorithm 2 consider the influence of sample concentration and instrument detection characteristics on data uncertainty. After many different combinations, suitable calculation methods for various heavy metals are obtained, as shown in Table 2.

**Table 1.** Comparison of two uncertainty algorithms

| Number | Algorithm | Related Parameters |
|--------|-----------|---------------------|
| 1      | \( U_{nc} = \sqrt{(\theta \times C)^2 + \text{MDL}^2} \) | Unc is the uncertainty; \( \theta \) is standard deviation; \( C \) is the measured heavy metal content; \( \text{MDL} \) is the detection limit |
| 2      | Unc=0.1C+\text{MDL}/3 | \( C \) is the measured heavy metal content; \( \text{MDL} \) is the detection limit |

**Table 2.** Appropriate uncertainty algorithms for eight heavy metals

| Component | Suitable algorithm |
|-----------|--------------------|
| As Zn Cu  | 1                  |
| Pb Cd Cr Ni Hg | 2                |

### 2.4 Evaluation Method of Heavy Metal Pollution

#### 2.4.1 Relative pollution assessment method

The pollution load index method is a combination of a single factor pollution index and a pollution load index \(^{[1-2]}\), which can reflect the relative pollution level of a single factor and all factors relative to the background value level (G, P et al. 2012). The calculation method is as follows.

\[ CF_i = \frac{C_i}{C_0} \quad (i = 1, 2, \ldots, n) \]

\( \text{PLI} = \sqrt[3]{CF_1 \times CF_2 \times \ldots \times CF_n} \)  

(3)

In the formula, \( C_i \) is the measured value of heavy metal \( i \) (mg/kg); \( C_0 \) is the background value of heavy metal \( i \). The background values of eight heavy metal used in this study are the background values of purple soil in Chongqing published by Yin Qihou et al. In 1985 \(^{[3]}\), \( n \) is the number of heavy metals evaluated, \( CF_i \) is the pollution index of heavy metals \( i \), and PLI is the pollution load index.

### 3. Results and Discussion

#### 3.1 Evaluation Results of Pollution Load Index Method
The pollution indexes of different heavy metals in farmland soil are \( CD > Pb > Hg > Zn > Cr > Ni > Cu > As \), among which the average pollution indexes of \( Cd \) is 2.93, reaching the level of moderate pollution; the pollution indexes of \( Pb, Hg, Zn \) and \( Cr \) are 1.13, 1.12, 1.11 and 1.10 respectively, belonging to light pollution; and the average pollution indexes of \( Ni, Cu \) and \( As \) (some points are seriously over the standard) are not more than 1, showing no pollution.

3.2 Analysis of Source Apportionment Results of PMF Model

In this study, the average value of \( Hg \) in the surface layer of farmland soil is lower than the background value in the whole study area. However, there are obvious high value areas and low value areas in different areas, while the high value areas are near a power plant in Luohuang town. At the same time, the content of \( Cd \) in this area is also significantly increased, and studies have shown that the large-scale use of coal may cause \( Cd \) pollution (Wen Mingzhong [4], 2010). Therefore, the main factors that lead to the increase of \( Hg \) and \( Cd \) content in the soil are the volatilization of \( Hg \) and \( Cd \) caused by coal combustion or the atmospheric subsidence caused by the enrichment in fly ash. So factor 1 is interpreted as the source of atmospheric deposition.

\( Pb \) and \( Zn \) make great contribution to factor 2, which are 43% and 41% respectively. This research area is the main industrial agglomeration area in Chongqing, which gathers a large number of polluting enterprises. Waste will inevitably be generated in the production process of enterprises, and the "three wastes" and other substances discharged will enter the surrounding environment. The spatial distribution of heavy metals in this study shows that the spatial distribution of heavy metals \( Zn \) and \( Cu \) is similar. The areas with serious pollution are mainly concentrated around zinc refineries, fertilizer plants and other enterprises. It can be seen that the sources of heavy metals \( Zn \) and \( Cu \) are closely related to the industrial production process. The coefficients of variation of \( Zn \) and \( Cu \) are 20% and 35% respectively, indicating that their distribution is obviously affected by human activities. There is a positive correlation between \( Zn \) and \( Cu \) \((P < 0.01)\), which further indicates that the accumulation of heavy metals as \( Zn \) and \( Pb \) in the surface soil in the farmland in the study area has a similar source, that is mainly from industrial emissions \([5-6]\). Therefore, factor 2 is an industrial source.

The contribution rate of factor 3 is from \( Ni, Cr, as, Cu, \) and \( Zn \), to \( Pb \), and the contents of \( Ni, Cr \) and \( Cu \) are lower than the background value and the contribution rate is more than 50%. As a whole, the spatial distribution of \( Ni \) and \( Cr \) in the surface soil in the farmland in the study area is relatively uniform, and the coefficient of variation \( Ni \) (19%), \( Cr \) (23%) and as (25.9%) are relatively small, which further shows that the three elements of \( Ni, Cr \) and \( As \) are less affected by human activities in the whole study area. It can be seen that the elements of \( Ni, Cr \) and \( As \) are mainly from natural sources, which are affected by soil parent material \([7-8]\). Therefore, factor 3 is determined to be the natural parent material source.

The data of this study shows that the average value of \( Cd \) is 3 times higher than the background value of \( Cd \). The high value of \( Cd \) is mainly concentrated in the agricultural area without industrial and mining enterprises except a small amount near the industrial park. Therefore, the main source of soil \( Cd \) is excluded for industrial sources and atmospheric deposition. Studies have shown that long-term fertilization may cause \( Cd \) accumulation in farmland soil in this area (Wang Mei [9], 2014), and the increase of \( Cd \) content in soil is relatively large compared with the background value, so the source of this factor excludes industrial sources and natural factors, and its pollution source can be interpreted as agricultural sources.

The four emission factors are analyzed by arithmetic weighted average. And the results indicate that the contribution rate of atmospheric deposition source, industrial source, natural parent material source and agricultural source is 19.7%, 27%, 35.5% and 17.8% respectively.

4. Conclusion

Four contribution sources are analyzed with the help of PMF model. And the contribution rate of atmospheric deposition source, industrial source, natural parent material source and agricultural source is 19.7%, 27%, 35.5% and 17.8% respectively. It shows that human factors (64.5%) such as industrial
and agricultural production are the main contributors to the accumulation of heavy metals in the soil in the study area. Therefore, when formulating regional development policies, the impact of industrial pollution on heavy metals in farmland surface soil should be considered [10]. In addition, it is necessary to restrict the development of high-energy-consuming and high-pollution enterprises, and reduce pollution sources and pollutant emissions. And it needs to use chemical fertilizers and pesticides reasonably to ensure the quality of agricultural soil.

Acknowledgments
Fund Project: source analysis and prevention and control of heavy metal Cd pollution in farmland of Hechuan City (cstc2018jxj20015)

References
[1] Xue Xiaojun, Niu Haibo, Yang Yang, etc. 2019. *Wet Zinc Smelting and Purifying Copper Slag Resource Recycling Technology.* Non-ferrous Equipment, (3): 19-21 + 4.
[2] Li Jiao, Teng Yanguo, Wu Jin, Jiang Jinyuan, Huang Ying. *Source Analysis of Soil Heavy Metals in the Middle and Upper Reaches of Le’an River Based on PMF Model and Geostatistics.* Environmental Science Research, 2019, 32 (06): 984-992.
[3] Yin Qihou, Xu Maoqi, Chen Nian, Du Zirong, Mu Shusen, Wang Gang. *Geochemical Background Values of Some Trace Elements in Purple Soil in Chongqing Area.* Journal of Southwest Normal University (Natural Science Edition), 1985 (03): 45-53.
[4] Wen Mingzhong, Xu Wendong. *Migration and Release of Harmful Trace Elements in Coal-fired Power Plants.* Journal of China Coal Society, 2010 (09): 108-113.
[5] Wang Mei, Li Shutian. *The Heavy Metal Content of Fertilizers and the Effect of Fertilization on the Accumulation of Heavy Metals in Soil and Crops.* Journal of Plant Nutrition and Fertilizer, 2014, 000 (002): 466-480.
[6] Jiang Yuling, Ruan Xinling, Ma Jianhua. *Characteristics and Classification Management of Heavy Metal Pollution in Sewage Irrigation Farmland near a Battery Plant in Xinxiang City.* Journal of Environmental Science, 2020, 40 (02): 645-654.
[7] Chen Zhifan, Hua Yanxu, Xu Wei, Pei Jincheng. *Analysis of Pollution Sources of Heavy Metals in Suburban Farmland Based on PMF Model.* Journal of Environmental Science, 2020, 40 (01): 276-283.
[8] Xu Yan, Wang Shuguang, Guo Zhen, Lu Yangjie. *Application of Principal Component Analysis in the Analysis of Farmland Soil Pollution in Mining Areas.* Environmental Science and Technology, 2019, 42 (S2): 9-13.
[9] Zeng Xiaoduo, Wang Xiangqin, Tu Xinhong, Zou Huaxu, Liu Chuanping, Li Fangbai, Wu Qitang. *Research Progress of Pollution Control Technology of Heavy Metal in Farmland Soil.* Journal of Eco-Environment, 2019, 28 (09): 1900-1906.