1. Assessment methods of soil heavy metal

The calculation formula of Contamination factor (CF) is as follows [1]:

$$CF=\frac{C_m(sample)}{C_m(background)} \quad (1)$$

where C_m is the measured values of individual heavy metals; Class 1, CF < 1 indicates low contamination; Class 2, 1 ≤ CF < 3 indicates moderate contamination; Class 3, 3 ≤ CF < 6 indicates considerable contamination; Class 4, CF > 6 indicates very high contamination.

The calculation formula of Pollution load index (PLI) is as follows [1]:

$$PLI_n=\sqrt[i]{CF_1 \times CF_2 \times CF_3 \times \cdots \times CF_i} \quad (2)$$

$$PLI_{zone}=\sqrt[n]{PLI_1 \times PLI_2 \times PLI_3 \times \cdots \times PLI_n} \quad (3)$$

where PLI_n is the pollution load index at the nth sampling site, PLI_{zone} is the pollution load index across the study area. PLI > 1 suggests that pollution exists; otherwise, there is no pollution.

The calculation formula of Nemerow pollution index (P_N) is as follows [2]:

$$P_N=\sqrt{\frac{{CF_{max}}^2+{CF_{mean}}^2}{2}} \quad (4)$$

where CF_{max} and CF_{mean} are the maximum and mean values, respectively. The pollution level of soil can be classified as: safe ($P_N \leq 0.7$), precaution ($0.7 < P_N \leq 1.0$), slightly polluted ($1.0 < P_N \leq 2.0$), moderately polluted ($2.0 < P_N \leq 3.0$), and seriously polluted ($P_N >$ 3.0).

The calculation formula of geo-accumulation index (I_{geo}) is as follows [1]:

$$I_{geo} = log_2\left(\frac{C_i}{1.5C_b}\right) \quad (5)$$

where C_i and C_b are the measured and background values of individual heavy metals, respectively. and 1.5 is the natural enrichment coefficient of heavy metals in the soil (Loska, Wiechula and Korus 2004). The grading criteria of the I_{geo} was shown in Table S2.

The calculation formula of enrichment factor (EF) is as follows [3]:

$$EF=\frac{(C_i/C_{ref})_{sample}}{(C_i/C_{ref})_{background}} \quad (6)$$

where $(C_i/C_{ref})_{sample}$ represents the heavy metal to reference element concentration ratio for soil sample, and $(C_i/C_{ref})_{background}$ represents the heavy metal to reference element concentration ratio for background value. Mn is selected as the reference metal. The grading criteria of the EF was shown in Table S2.

The improved Nemerow index (INI) method was applied to determine the integrated pollution level of soil heavy metals [2–4]. Its equation is as follows:

$$INI = \sqrt{\frac{{I_{geo_{max}}}^2+{I_{geo_{avg}}}^2}{2}} \quad (7)$$

where $I_{geo_{max}}$ and $I_{geo_{avg}}$ are the maximum and average values of the I_{geo} for sample points, respectively. The grade classification criteria was shown in Table S2.

The risk assessment code (RAC) expressed the ecological risk of soil heavy metals through the proportion of their available fraction to the total content. Its arithmetic equation is as follows [5]:

$$\text{RAC} = \frac{C_{ava}}{C_i} \quad (8)$$

where RAC is the risk code value, C_{ava} is the available state of soil heavy metals, and C_i is the content of heavy metal in soil, the corresponding class classification was shown in Table S3.

The potential ecological risk (RI) takes into account not only the concentration of heavy metal, but also the toxicity of heavy metals and other factors, and can obtain a more comprehensive risk evaluation result [5]. Its arithmetic equation is as follows:

$$E_r^i = T_r^i \times \frac{C_i}{C_b} \quad (9)$$

$$\text{RI} = \sum_{i=1}^{n} E_r^i \quad (10)$$

where E_r^i is the ecological risk value of individual heavy metals, T_r^i is the biotoxicity factor value of individual heavy metals, the values of T_r^i for V, Cr, Co, Cu, Zn, Cd, Pb, Hg and As are 2, 2, 5, 5, 1, 30, 5, 40 and 10, respectively [6]. C_i and C_b are the concentration and background values of a single heavy metal, respectively. RI is the potential ecological risk index. The evaluation criteria were shown in Table S3.

The human health risk assessment (HHRA) was presented as follows [3]：

$$\text{ADD}_{\text{ingest}} = \frac{C \times R_{ingest} \times EF \times ED}{BW \times AT} \times 10^{-6} \quad (11)$$

$$\text{ADD}_{\text{dermal}} = \frac{C \times SA \times SL \times ABF \times EF \times ED}{BW \times AT} \times 10^{-6} \quad (12)$$

$$\text{ADD}_{\text{inhal}} = \frac{C \times R_{inhal} \times EF \times ED}{PEF \times BW \times AT} \quad (13)$$

$$\text{HI} = \sum HQ = \sum \frac{ADD_{ij}}{RfD_{ij}} \quad (14)$$

$$\text{TCR} = \sum CR = \sum ADD_{ij} \times SF_{ij} \quad (15)$$

Where $\text{ADD}_{\text{ingest}}$, $\text{ADD}_{\text{dermal}}$, and $\text{ADD}_{\text{inhal}}$ are the average daily dose of soil by ingestion, dermal and inhalation absorption, respectively [7, 8]. C is the concentration of heavy metals in soil, HI and TCR are hazard index and total carcinogenic risk, respectively. HQ and CR are hazard quotient and carcinogenic risk, respectively. The values and meanings of other parameters were shown in Table S4 and S5, and the hierarchy was displayed in Table S6 [7, 8].

2. Positive matrix factorization

Positive matrix factorization (PMF) is a source apportionment method recommended by the U.S. Environmental Protection Agency (U.S. EPA) and was first proposed by Paatero and Tapper [9, 10]. PMF model divides the sample data matrix into two matrices based on chemical mass balance principle: factor contribution (G) and factor profile (F), as shown in Equation:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \tag{16}$$

where x_{ij} is the concentration matrix of element *j* in sample *i*; g_{ik} is the contribution of factor *k* in sample *i*; f_{kj} is the matrix of chemical compositions of factor; e_{ij} is the matrix of residual.

Factor contributions and profiles are derived by the PMF model minimizing the objective function Q defined as follows:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{e_{ij}}{u_{ij}} \right)^2 \tag{17}$$

where *Q* is the sum of the squares of the ratios of the residuals and uncertainties (u_{ij}), The uncertainty data given by the authors or calculated from the formula.

Reference

[1] Guan, J.; Wang, J.; Pan, H.; Yang, C.; Qu, J.; Lu, N.; Yuan, X. Heavy metals in Yinma River sediment in a major Phaeozems zone, Northeast China: Distribution, chemical fraction, contamination assessment and source apportionment. *Sci Rep* **2018**, *8(1)*, 1–11. https://doi.org/10.1038/s41598-018-30197-z.

[2] Liu, H.; Zhang, Y.; Yang, J.; Wang, H.; Li, Y.; Shi, Y.; Hu, W. Quantitative source apportionment, risk assessment and distribution of heavy metals in agricultural soils from southern Shandong Peninsula of China. Sci Total Environ 2021, 767, 144879. https://doi.org/10.1016/j.scitotenv.2020.144879.

[3] Huang, J.; Wu, Y.; Sun, J.; Li, X.; Geng, X.; Zhao, M.; Fan, Z. Health risk assessment of heavy metal (loid) s in park soils of the largest megacity in China by using Monte Carlo simulation coupled with Positive matrix factorization model. J Hazard Mater 2021, 415, 125629. https://doi.org/10.1016/j.jhazmat.2021.125629.

[4] Santos-Francés, F.; Martínez-Graña, A.; Alonso, R. P.; García Sánchez, A. Geochemical background and baseline values determination and spatial distribution of heavy metal pollution in soils of the Andes mountain range (Cajamarca-Huancavelica, Peru). *Int J Environ Res Public Health* **2017**, *14(8)*, 859. https://doi.org/10.3390/ijerph14080859.

[5] Zhang, G.; Bai, J.; Xiao, R.; Zhao, Q.; Jia, J.; Cui, B.; Liu, X. Heavy metal fractions and ecological risk assessment in sediments from urban, rural and reclamation-affected rivers of the Pearl River Estuary, China. Chemosphere 2017, 184, 278–288. https://doi.org/10.1016/j.chemosphere.2017.05.155.

[6] Liu, J.; Liu, Y.; Liu, Y.; Liu, Z.; Zhang, A. Quantitative contributions of the major sources of heavy metals in soils to ecosystem and human health risks: A case study of Yulin, China. Ecotoxicol Environ Saf 2018, 164, 261–269. https://doi.org/10.1016/j.ecoenv.2018.08.030.

[7] Xiao, R.; Guo, D.; Ali, A.; Mi, S.; Liu, T.; Ren, C.; Zhang, Z. Accumulation, ecological-health risks assessment, and source apportionment of heavy metals in paddy soils: A case study in Hanzhong, Shaanxi, China. Environ Pollut 2019, 248, 349–357. https://doi.org/10.1016/j.envpol.2019.02.045.

[8] Huang, Y.N.; Dang, F.; Li, M.; Zhou, D.M.; Song, Y.; Wang, J.B. Environmental and human health risks from metal exposures nearby a Pb-Zn-Ag mine, China. Sci Total Environ 2020, 698, 134326. https://doi.org/10.1016/j.scitotenv.2019.134326.

[9] Wang, Y.; Guo, G.; Zhang, D.; Lei, M. An integrated method for source apportionment of heavy metal (loid) s in agricultural soils and model uncertainty analysis. Environ Pollut 2021, 276, 116666. https://doi.org/10.1016/j.envpol.2021.116666.

[10] Guan, Q.; Wang, F.; Xu, C.; Pan, N.; Lin, J.; Zhao, R.; Luo, H. Source apportionment of heavy metals in agricultural soil based on PMF: A case study in Hexi Corridor, northwest China. Chemosphere 2018, 193, 189–197. https://doi.org/10.1016/j.chemosphere.2017.10.151.

Table S1. The information on the reclamation years of sample sites.

Sample number	Reclamation years	Longitude (°E)	Latitude (°N)
1	60	124°59′24″	45°0′57″
2	60	124°59′24″	45°0′57″
3	60	124°57′15″	45°0′29″
4	60	124°52′9″	45°0′57″
5	60	124°52′9″	45°0′57″
6	50	124°47′31″	45°0′21″
7	50	124°47′31″	45°0′21″
8	50	124°43′25″	44°59′4″
9	50	124°43′25″	44°59′4″
10	30	124°39′50″	44°58′30″
11	30	124°39′50″	44°58′30″
12	30	124°39′50″	44°58′30″
13	30	124°39′50″	44°58′30″
14	30	124°37′52″	44°57′3″
15	30	124°37′52″	44°57′3″
16	30	124°37′52″	44°57′3″
17	30	124°37′52″	44°57′3″
19	30	124°38′7″	45°3′32″
20	20	124°42′32″	45°4′53″
21	20	124°42′32″	45°4′53″
22	20	124°39′43″	45°9′15″
23	20	124°39′43″	45°9′15″
24	20	124°38′19″	45°9′33″

Table S2. The grade classification criteria of geo-accumulation index and improved Nemerow index.

Classification level		Value	Reference
Geo-accumulation index			
class 0	Unpolluted	≤ 0	Guan et al. 2018
class 1	Unpolluted to moderately polluted	$0 < I_{geo} \leq 1$	
class 2	Moderately polluted	$1 < I_{geo} \leq 2$	
class 3	Moderately to heavily polluted	$2 < I_{geo} \leq 3$	
class 4	Heavily polluted	$3 < I_{geo} \leq 4$	
class 5	Heavily to extremely polluted	$4 < I_{geo} \leq 5$	
class 6	Extremely polluted	$I_{geo} > 5$	
Enrichment factor			
class 1	Minimal enrichment	EF < 2	Huang et al. 2021
class 2	Moderate enrichment	2 ≤ EF < 5	
class 3	Significant enrichment	5 ≤ EF < 20	
class 4	Very high enrichment	20 ≤ EF < 40	
class 5	Extremely high enrichment	EF ≥ 40	
Improved Nemerow index			
class 0	Uncontaminated	INI < 0.5	Santos-Francés et al. 2017
class 1	Uncontaminated to moderately contaminated	0.5 ≤ INI < 1	
class 2	Moderately contaminated	1 ≤ INI < 2	
class 3	Moderately to heavily contaminated	2 ≤ INI < 3	

class 4	Heavily contaminated	3 ≤ INI < 4
class 5	Heavily to extremely contaminated	4 ≤ INI < 5
class 6	Extremely contaminated	INI ≥ 5

Table S3. The evaluation criteria of the potential ecological risk index and risk assessment code.

Classification level	**Value**	**Reference**
The ecological risk index of individual heavy metals		
Low risk	E_r^i < 40	Zhang et al. 2017
Moderate risk	40 ≤ E_r^i < 80	
Considerable risk	80 ≤ E_r^i < 160	
High risk	160 ≤ E_r^i < 320	
Extremely high risk	E_r^i ≥ 320	
The potential ecological risk index		
Low risk	RI <150	Zhang et al. 2017
Moderate risk	150 ≤ RI < 300	
Considerable risk	300 ≤ RI < 600	
High risk	600 ≤ RI < 1200	
Extremely high risk	RI ≥ 1200	
The risk assessment code		
No risk	RAC ≤ 1	Zhang et al. 2017
Light risk	1 < RAC ≤ 10	
Medium risk	10 < RAC ≤ 30	
High risk	30 < RAC ≤ 50	
Very high risk	50 > RAC	

Table S4. The meanings and values of the parameters in the human health risk assessment.

Parameters	Description	Units	Values		Reference
			children	adults	
R_{ingest}	ingestion rate of soil	mg/day	200	100	USEPA 2011
EF	exposure frequency	day/year	350	350	Jiang et al. 2020
ED	exposure duration	year	6	26	Jiang et al. 2020
BW	average body weight	kg	29.3	62.57	Huang et al. 2020
AT	average time of exposure to contaminated soils	day	365×*ED* (non-carcinogenic)		Huang et al. 2021
			365×70 (carcinogenic)		Huang et al. 2021
SA	exposed skin area	m^2	0.23	0.57	Jiang et al. 2020
SL	skin adherence factor	mg/cm^2	0.2	0.07	USEPA 2011
ABF	dermal adsorption factor	unitless	0.001		Jiang et al. 2020
R_{inhal}	Inhalation rate	m^3/day	7.5	16.1	MEPPRC 2014
PEF	particle emission factor	m^3/kg	1.36×10^9		USEPA 2011

Table S5. Corresponding reference dose (*RfD*) and slope factors (*SF*) values of metals by different exposure pathways used in the human health risk assessment.

Element	*RfD* (mg/(kgd))			*SF*((mgd)/mg)		
	ingestion	inhalation	dermal contact	ingestion	inhalation	dermal contact
V	7.00E-03		7.00E-05			
Cr	3.00E-03	2.86E-05	6.00E-05	5.00E-01	4.20E+01	2.00E+01
Co	2.00E-02	5.71E-06	1.60E-02		9.80E+00	
Cu	4.00E-02	4.02E-02	1.20E-02			
Zn	3.00E-01	3.00E-01	6.00E-02			
Cd	1.00E-03	1.00E-05	1.00E-05	5.10E-01	6.30E+00	2.00E+01
Pb	3.50E-03	3.50E-03	5.25E-04	8.50E-03		
Hg	3.00E-04	8.57E-05	2.10E-05			
As	3.00E-04	1.23E-04	1.23E-04	1.50E+00	1.51E+01	3.66E+00

Reference: Huang et al. 2021; Xiao et al. 2019; Huang et al. 2020.

Table S6. The classification of human health risk levels.

Classification level	Value	Reference
The non-carcinogenic risk		
There is potential adverse health effect	HQ or HI >1	Huang et al. 2021
There is not potential adverse health effect	HQ or HI ≤1	
The carcinogenic risk		
Negligible cancer risk	TCR < 10^{-6}	Huang et al. 2021
Acceptable cancer risk	10^{-6} ≤ TCR < 10^{-4}	
Significant cancer risk	10^{-4} ≤ TCR	

Table S7. Class distribution for pollution assessment of heavy metals by Pollution load index (PLI).

Pollution degree	No pollution	pollution
PLI_n	0	100%
PLI_{zone}		100%

Table S8. Class distribution for pollution assessment of heavy metals by Nemerow index (P_N).

Pollution degree	safe	precaution	slightly polluted	moderately polluted	seriously polluted
P_N	0	0	0	0	100%

Table S9. Class distribution for pollution assessment of heavy metals by improved Nemerow index (INI).

Classification level	Class distribution
Improved Nemerow index	
The mean value of INI	1.55
Uncontaminated to moderately contaminated	5.71%
Moderately contaminated	91.43%
Moderately to heavily contaminated	2.86%

Table S10. Class distribution for pollution assessment of heavy metals by potential ecological risk index (RI).

Classification level	Class distribution
The potential ecological risk index	
The mean value of RI	408.28
Moderate risk	22.86%
Considerable risk	74.28%
High risk	2.86%

Table S11. Rotated component matrix for heavy metals in soil.

Heavy metal	PC1	PC2	PC3	PC4
V	**0.77**	0.42	0.27	0.02
Cr	**0.91**	0.18	-0.08	0.00
Co	0.68	0.45	0.49	0.13
Cu	0.54	**0.76**	0.09	0.02
Zn	**0.80**	0.38	0.40	0.09
Cd	0.04	0.09	0.04	**0.99**
Pb	0.50	0.63	0.47	-0.02
Hg	0.10	0.16	**0.96**	0.04
As	0.22	**0.87**	0.17	0.15
Percentage of variance (%)	34.63	25.67	18.52	11.45
Percentage of cumulative variance (%)	34.63	60.30	78.82	90.27

note: PC1, 2, 3 and 4 refer to principal component 1, 2, 3 and 4.

Table S12. List of Abbreviations.

Abbreviation	Definition	Unit
HMs	heavy metals	mg/kg
CF	contamination factor	-
I_{geo}	geo-accumulation index	-
EF	enrichment factor	-
PLI	pollution load index	-
PN	Nemerow pollution index	-
INI	improved Nemerow index	-
RAC	risk assessment code	%
RI	potential ecological risk index	-
HHRA	human health risk assessment	-
CR	carcinogenic risk	-
NCR	non-carcinogenic risk	-
HI	hazard index	-
HQ	hazard quotient	-
TCR	total carcinogenic risk	-
SD	standard deviation	-
CV	coefficient of variation	-
BV	background values	mg/kg
RSV	the screening value of agricultural land	mg/kg
N_{exceed}	the rate of the number for soil samples exceeding the background value to the total samples number	%