



# Source apportionment and risk assessment of heavy metals in urban soils from a central China city by using positive matrix factorization model coupled with Monte Carlo simulation

Lianfeng Shen<sup>1</sup> · Jin Zeng<sup>1</sup> · Erhui Guo<sup>1</sup> · Yongxia Cheng<sup>2</sup> · Limei Li<sup>1</sup> · Cunyang Niu<sup>1</sup> · Yuan Yuan<sup>1</sup> · Li Ma<sup>1</sup>

Accepted: 2 August 2022 / Published online: 17 August 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

The concentrations of As, Cd, Co, Cu, Mn, Ni, Pb and Zn in 78 urban soil samples from Zhengzhou, Henan Province, China were determined in the study. The mean values of the eight heavy metals were lower than the natural background reference values of soils from Henan Province, with the exception of Cd and Pb. However, the peak values for six species of the elements exceed the corresponding reference values, indicating enrichment of these metals in urban soils in Zhengzhou City, especially for Cd and Pb. The spatial distribution patterns showed that levels of As, Co, Mn and Ni decreased gently from northwest to southeast without hot-spot area, whereas Cd, Cu, Pb and Zn were with significant spatial variability. PMF model coupled with Pearson's correlation analysis was performed to apportion and quantify possible sources of the eight elements. Four Factors (natural source, industrial production, agricultural activities and traffic discharges) were distinguished and their respective contribution rates were 18.73%, 33.22%, 23.91% and 24.14%, respectively. Results of uncertainty analysis confirmed that the all four source Factors based on the PMF model were reliable. Furthermore, the probability distribution of health risk assessed by means of Monte Carlo simulation revealed that non-carcinogenic risks associated with heavy metals were acceptable for both adults and children, while the total carcinogenic risk values remained relatively high. Additionally, the children were more susceptible to health risks in comparison to adults, and the ingestion way was considered to be the major exposure pathway. The results of source-oriented health risk demonstrated that the agricultural source was the major contributor to health risks, which was characterized by As, Ni and Co.

**Keywords** Pollution assessment · Potential ecological risk · Zhengzhou urban area · Source identification · Probabilistic health risk

## 1 Introduction

Urban soils are regarded as a crucial component of the city ecological system which closely related to the general life quality of residents (Pecina et al. 2021). However, with the accelerated urbanization and industrial expansion in and

around cities, frequent anthropogenic disturbances have led to the accumulation of various pollutants in urban soils. Among numerous types of pollutants, heavy metal contaminants are of great concern due to their biotoxicity and irreversibility, and they are considered to be important indicators for monitoring the impact of anthropogenic processes on soil environmental quality (Tong et al. 2020).

The sources of heavy metals in the soil can be either natural or anthropogenic or both (Siddiqui et al. 2020). Urban soils differ significantly from other natural ones in that they are more susceptible to human activity. Therefore, anthropogenic sources are seen as the primary cause of the increased pollution elements in urban soils, for instance, vehicle emissions, municipal and industrial wastes disposal, mining and smelting activities, construction waste

✉ Li Ma  
malizhk@henau.edu.cn

<sup>1</sup> College of Forestry, Henan Agricultural University, Zhengzhou 450002, China

<sup>2</sup> Key Laboratory for Monitor and Remediation of Heavy Metal Polluted Soils of Henan Province, Jiyuan 459000, China

and domestic waste, fossil fuel combustion, huge usage of agrochemicals (Ma et al. 2016; Men et al. 2018; Hou et al. 2019; Faisal et al. 2021). It is reported that over five million sites were contaminated by heavy metals worldwide (Sun et al. 2019). In view of the non-degradable properties, heavy metals tend to accumulate in organisms, and they may damage the nervous, lungs, liver, kidneys skeletal, enzymatic, endocrine, and immune systems promoting several disease conditions and even lead to cancer ultimately (Jaishankar et al. 2014; Li et al. 2014; Vareda et al. 2019). Heavy metals can do harm to humans in three ways: ingestion, inhalation, and dermal contact (Safiur Rahman et al. 2019). Previous studies have showed that ingestion is the as the main exposure route for risk of humans, furthermore, children were more susceptible health threats from toxic elements than adults (Tong et al. 2019, 2020; Zhang et al. 2019; Huang et al. 2021; Siddiqui et al. 2021).

Since the origins of elements in urban soils were complex and varied, source identification and distinction of heavy metals was regarded as precondition of preventing and controlling pollution (Wu et al. 2018). The positive matrix factor (PMF) model cannot only allow for non-negative constraints on both the loadings and contribution, but can also handle some missing data (Li et al. 2020). In consequence, the PMF model has been used previously to apportion the source allocation of atmospheric, water, sediment and soil pollutants (Zhang et al. 2018, 2020; Liu et al. 2019b; Srivastava et al. 2021). Nonetheless, the approach relies predominantly on the researcher's analysis for the local background parameter values, and its accuracy and application were constrained (Huang et al. 2021). Consequently, a comprehensive research method based on PMF in conjunction with correlation coefficients was employed to ensure the reliability of source assignment results in the study.

Meanwhile, the health risk levels varied by element type and population group, and the contribution of each pollution source to health risk accounts for different shares. To accurately control the health risks related to heavy metals in soil, there is a need for conducting the health risk assessment and identify highlight risky elements. In term of health risk assessments, previous studies tended to depend upon traditional models with fixed parameters, resulting in overestimation or underestimation results (Chen et al. 2019; Zeng et al. 2019; Qiao et al. 2020). Fortunately, Monte Carlo simulation technique, which take both the variability of specific exposure parameters and uncertainty of heavy metal concentrations into account, can minimize the uncertainties in the evaluation of health risks. Given this advantage of the Monte Carlo simulation approaches, it

has been shown to be one of the most powerful method for probabilistic risk assessment (Ma et al. 2017; Wang et al. 2020b). Therefore, it was employed to evaluate health risks resulting from soil elements and to identify preferent heavy metals for risk control in the study.

Urban soils have been regarded as a good tool for diagnosing environmental pollutions (Awasthi et al. 2016; Liu et al. 2009). The spatial distributions, pollution level analysis, source identification and concentration-oriented health risk caused by toxic elements in urban soils have been studied a lot (Wang et al. 2017, 2019; Men et al. 2018; Sun et al. 2019). However, very few studies explored source apportionment and source-oriented health risk of heavy metals in urban soils, especially that in large megacity. As the core city of central plains city group, the soil pollution characteristics and the related health risks to human beings in the urban areas of Zhengzhou remains to be lack of study, in spite of the fact that many people may be involved. On account of this, the study focused on the contamination characteristics and probabilistic health risk of pollution elements in urban soils from Zhengzhou. The core objectives include (1) to investigate pollution characteristics and the potential ecological risk of eight heavy metals in urban soils, (2) to identify and quantify the source apportionment of targeted pollutants, (3) to evaluate concentration-oriented and source-oriented health risk for both adult and children exposed to heavy metals.

## 2 Experimental

### 2.1 Site description

Zhengzhou city (112°42'–114°14' E, 34°16'–34°58' N) which locates in the geographical center of China is the capital of Henan Province, and is an important transportation junction, as well as a national logistics and population catchment area. The city covers an area of 7,446 km<sup>2</sup>, and it contains 12.6 million residents according to the seventh national census. The climate of Zhengzhou is characterized as northern temperate continental monsoon climate with an average rainfall of 633.1 mm per year. In the past decades, the predominant geological location, policies, cultural conditions played key roles in attracting more people to visit or settle in the city. While the huge population drives economic development, and it also contributes directly or indirectly to the formation of massive amounts of pollutants (Wang et al. 2020a). In Zhengzhou, there are ring roads surrounding the Erqi Tower up to various distances. The areas within 4th ring road have the

higher population density and vehicle stream, and it's the regions of concern for the study.

## 2.2 Sample collection and preparation

In this study, the area within 4th ring road of Zhengzhou was divided into a sampling grid of 2.5 km × 2.5 km. A total of 78 urban soils were gathered in March 2019 (Fig. 1), and renovation soil sites were avoided. At each sampling location, no less than 5 subsamples were sampled using a stainless steel spatula within the radius of 50 m. Samples at each sub-sites were blended to obtain one composite sample and sealed in a self-sealing polyethylene pouch subsequently. After being naturally dried at room temperature, stone oak, construction debris, dry branches and fallen leaves were removed from the samples. The remaining were abraded and passed through a 0.15-mm nylon sieve, and then the homogeneous sieved samples were kept in polyethylene bags before analysis.

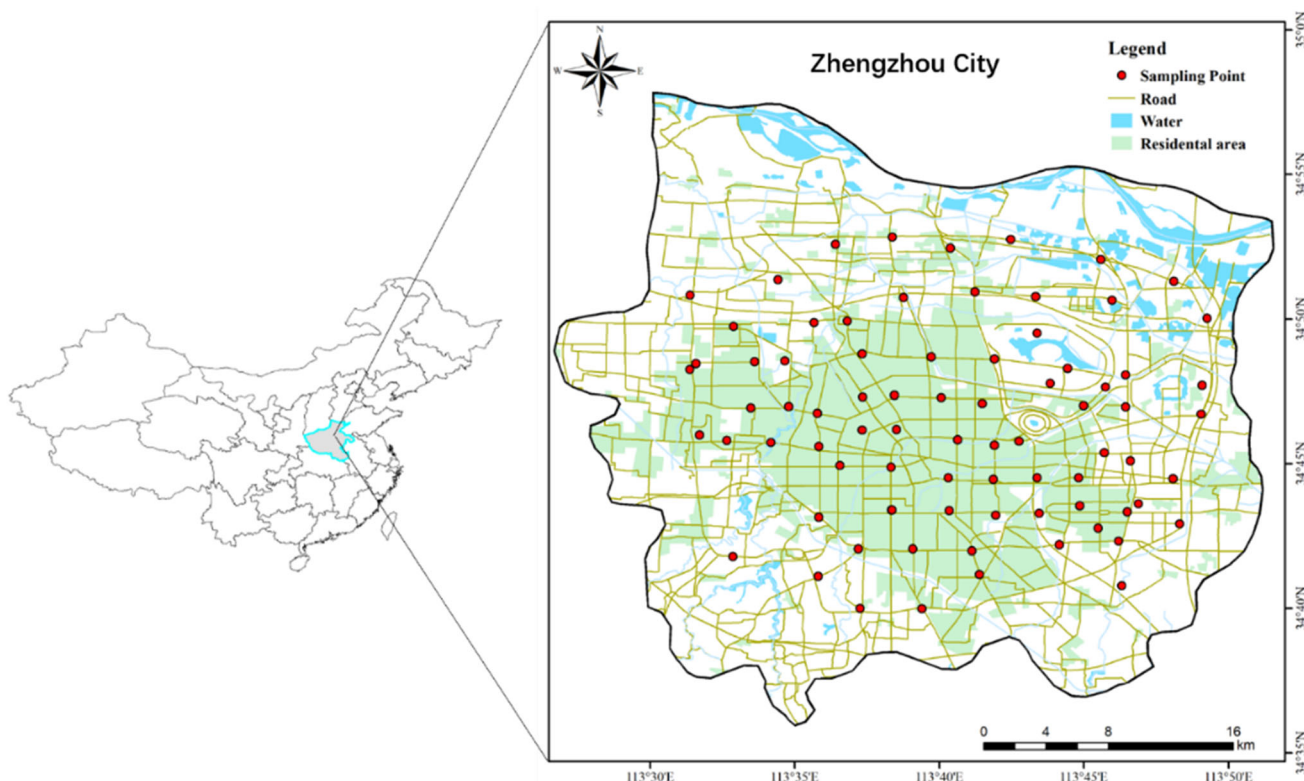
## 2.3 Heavy metal determination

The contents of As, Cd, Co, Cu, Mn, Ni, Pb and Zn were measured according to previously described methods with minor modification (Ma et al. 2016). A microwave digestion (Mars One, CEM Corporation, Matthews, USA) was

employed to extract target analytes from soil samples. Approximately 0.50 g sample was weighed and digested under pressure into a polytetrafluoroethylene (PTFE) vessel with 9 mL of concentrated HNO<sub>3</sub> (70%) and 3 mL of HCl (37%). On completion of the digestion procedure, the digestion solutions were heated to less than 5 mL in the acid detector. After adequate cooling, solutions were filtered and then replenished to 50 mL with deionized water (18.2 MΩ cm, ULUP-III-10 T-ZL, Sichuan Youpu Ultra-pure Technology Co., Sichuan China). The element concentrations were determined by ICP-MS (Agilent 7700x, Tokyo, Japan). A certified reference material (CRM) soil GBW07444 was applied to verify the digestion efficiency. Results showed that the recovery rates for all the heavy metals were around 85–105%.

## 2.4 Statistical analysis

The descriptive statistics of eight metals were performed on the strength of SPSS software (version 21.0). Pearson correlation matrix analysis was performed by Origin 2020b to quantify the relationship between each pair of targeted pollutants. Positive matrix factorization (PMF) model was carried out through the EPA PMF software (version 5.0) to allocate the hypothetical sources of pollutants. Given the great impact of outliers on PMF results, the dataset was



**Fig. 1** Sampling sites distribution of urban soils in Zhengzhou City

detected before performing PMF model, and the outliers were identified and eliminated on the basis of interquartile ranges and histograms (Guan et al. 2018; Huang et al. 2021). In addition, to prevent the underestimation or overestimation of risks because of utilizing constant parameters, Monte Carlo simulation was selected to keep the uncertainties of the health risk evaluation to a minimum (Jiang et al. 2021). Crystal Ball software (version 11.1.24) was applied to perform the Monte Carlo simulation by considering 10,000 iterations with 95% confidence level. The spatial distribution of the contaminants in study area was drawn by Geographic Information System (GIS) software. The geostatistical interpolation method of ordinary kriging was selected to predicted the heavy metal concentrations at unsampled area (Zhang et al. 2021).

### 2.5 Ordinary kriging interpolation method

Geostatistical analyses based on ArcGIS software were used to describe all spatial variations. In order to develop better predictions, ordinary kriging interpolation method was introduced to estimate the attributes values at non-sampled sites by combining observed values with weights linearly. The predicted value could be calculated as follows:

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \tag{1}$$

where  $Z^*(x_0)$  is the predicted value at an unknown location  $x_0$ ,  $Z(x_i)$  is the observed values at location  $x_i$ ,  $\lambda_i$  is the weight placed on  $Z(x_i)$ .

### 2.6 Pollution levels and potential ecological risk of heavy metals

Geo-accumulation index ( $I_{geo}$ ) was applied to assess the pollution level of individual heavy metal in soils, and it can be achieved by the following formula (Huang et al. 2021):

$$I_{geo} = \log_2\left(\frac{C_i}{1.5 \times B_i}\right) \tag{2}$$

where  $C_i$  is the determined content of metal,  $B_i$  represents the geochemical background value of individual element and 1.5 is the constant value allowing the analysis of fluctuations heavy metals based on natural processes.  $I_{geo}$  is unitless, and it can be divided into seven classes, as shown in Table S1.

Different from the  $I_{geo}$ , all the individual evaluation factors and highlights the contribution of the highest enrichment elements were taken into account in the Nemerow composite index ( $P_s$ ) method (Cui et al. 2021), which expressed as follows:

$$C_f^i = \frac{C_i}{B_i} \tag{3}$$

$$P_s = \sqrt{\frac{(C_{fmax}^i)^2 + (C_{fave}^i)^2}{2}} \tag{4}$$

where  $C_f^i$  is the individual pollution index of element  $i$ ,  $C_{fmax}^i$  represents the maximum value of  $C_f^i$ , and  $C_{fave}^i$  is the average of single pollution index of all the elements.  $P_s$  can be classified into five grades and listed in Table S1.

Ecological risk index ( $RI$ ) was selected to estimate the adverse effect caused by heavy metals on the urban ecosystem, which considers the toxicity and environmental response of different heavy metals based on the toxicity response (Liu et al. 2019a). The  $RI$  was defined as ((Mihankhah et al. 2020):

$$RI = \sum_{i=1}^n E_r^i = \sum_{i=1}^n T_r^i C_f^i = \sum_{i=1}^n [T_r^i (C_i/B_i)] \tag{5}$$

where  $T_r^i$  is the biological toxic response factor of each element. According to previous studies, the values of  $T_r^i$  for As, Cd, Co, Cu, Mn, Ni, Pb and Zn are 10, 30, 1, 5, 1, 2, 5 and 1, respectively (Hakanson 1980; Dehghan Madiseh et al. 2009).  $E_r^i$  represents the potential ecological risk factor for a given substance. The division of  $RI$  values was suggested as Table S1.

### 2.7 Source apportionment model and uncertainty analysis

#### 2.7.1 Source apportionment model

In the research, the PMF model was introduced for source allocation of heavy metals following the guidelines of the USEPA (USEPA 2014). The model divided the original matrix into two factor matrices  $g_{ik}$  and  $f_{jk}$  and a residual matrix  $e_{ij}$ . The basic equation can be formulated as follows:

$$X_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij} \tag{6}$$

where  $X_{ij}$  is the concentration of heavy metal  $j$  measured in sample  $i$ ,  $g_{ik}$  is the contribution of source  $k$  to sample  $i$  (mg/kg) and  $f_{kj}$  is the amount of element  $j$  from source  $k$ . The residual  $e_{ij}$  is calculated by adjusting  $g_{ik}$  and  $f_{kj}$  to minimum value of the objective function  $Q$ , which is defined as:

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{e_{ij}}{u_{ij}}\right)^2 \tag{7}$$

$$u_{ij} = \begin{cases} \frac{5}{6} \times MDLC \leq MDL \\ \sqrt{(ErrorFraction \times C)^2 + (0.5MDL)^2} C > MDL \end{cases} \tag{8}$$

where  $u_{ij}$  represents the uncertainty of heavy metal  $j$  in sample  $i$ , Error fraction is the measurement percentage, and  $MDL$  represents the method detection limit.

### 2.7.2 Methods of uncertainty analysis

The goodness of fit, bias and variability were considered to be the common method for testing the uncertainties of PMF model results (Men et al. 2019; Chai et al. 2021). In the research, fitting coefficients ( $r^2$ ) between soil heavy metals observed concentrations and predicted concentrations by PMF model were applied to evaluate goodness of fit, and it could be calculated based on formula below:

$$r^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{9}$$

where  $\hat{y}_i$  is the predicted concentration at site  $i$ ,  $y_i$  is the observed concentration at site  $i$ ,  $\bar{y}$  is the mean observed concentration and  $n$  is the number of sampling sites.

Bootstrap (BS) could be estimate effects from random errors and part effects of rotational ambiguity, and it was used to test bias and variability of PMF model results consequently (Paatero et al. 2014). The essence of the method is to generate replicate datasets from the original observations by re-sampling with replacement. The bootstrap samples ( $X^* = [x_1^*, x_2^*, \dots, x_n^*]$ ) were developed by sampling from empirical distribution function ( $F_n$ ) of original data set ( $X = [x_1, x_2, \dots, x_n]$ ). The statistical object ( $R^*(X^*, F_n)$ ) could be calculated from the following formula:

$$R^*(X^*, F_n) = \theta(F_n^*) - \theta(F_n) \tag{10}$$

where  $F_n^*$  represents the empirical distribution function,  $\theta(F_n^*)$  and  $\theta(F_n)$  are the estimation of parameter  $\theta$  of population distribution of  $X$ . The distribution of  $R$  and the parameter  $\theta$  could be obtained after repeating the sampling and calculation of  $R^*(X^*, F_n)$  for enough times.

The BS data sets were specified number of 100 in this study and then processed with PMF. For each BS run, the BS factors were compared with the base run factors. If no base factors correlate above the threshold for a given BS factor, that factor is considered “unmapped”. This process would be repeated for as many BS runs as the specifies. There could be instances when multiple BS factors from the same run may be mapped to the same base factor.

Based on the BS simulations, the coefficient of variation (CV) values were applied to reflect the deviation level of

heavy metals source apportionments, greater CV values suggest greater level of deviation (Liu et al. 2018). It could be obtained from standard deviation (SD) values divided by average values of the 200 iterations BS results.

### 2.8 Probabilistic health risk assessment

The hazard quotient (HQ) and carcinogenic risk (CR) were widely employed to quantify the potential health risk posed by chemical elements for both children and adults (Gu and Gao 2018; Islam et al. 2020; Lee et al. 2021). HQ and CR were calculated using the corresponding average daily dose (ADD), reference dose (RfD) and cancer slope factor (SF), respectively. Given that the risk generated by heavy metals in soil comes from ingestion ( $ADD_{ing}$ ), inhalation ( $ADD_{inh}$ ) and dermal contact ( $ADD_{derm}$ ), the ADD could be calculated by Monte Carlo simulation according the following Eqs. (11)–(13). The optimal probability distributions for concentration database of exposure heavy metals and exposure parameters used in the method were presented in Table S2 and Table S3.

$$ADD_{ing} = \frac{C \times R_{ing} \times EF \times ED}{BW \times AT} \times 10^{-6} \tag{11}$$

$$ADD_{inh} = \frac{C \times R_{inh} \times EF \times ED}{PEF \times BW \times AT} \tag{12}$$

$$ADD_{derm} = \frac{C \times SL \times SA \times ABF \times EF \times ED}{BW \times AT} \times 10^{-6} \tag{13}$$

The evaluation indicators of hazard index (HI) and total carcinogenic risk (TCR) for elements could be estimated by the following formulas (Han et al. 2020; Li et al. 2020):

$$HI = \sum_{i=1}^n HQ_i = \sum_{i=1}^n \left( \frac{ADD_{ing}}{RfD} + \frac{ADD_{inh}}{RfD} + \frac{ADD_{derm}}{RfD} \right)_i \tag{14}$$

$$TCR = \sum_{i=1}^n CR_i = \sum_{i=1}^n (ADD_{ing} \times SF + ADD_{inh} \times SF + ADD_{derm} \times SF)_i \tag{15}$$

When  $HQ < 1$  or  $HI < 1$ , adverse health effect of heavy metals to the population is unlikely. If  $HQ > 1$  or  $HI > 1$ , it indicates that harmful effects on health caused by soil heavy metals may occur. When CR is less than 1.00E–06, the carcinogenic risk of elements is deemed negligible. When CR between 1.00E–06 and 1.00E–04, indicates that the risk to human beings is within the acceptable range. When CR is ore than 1.00E–04, shows that the carcinogenic risk produced by pollutants is unacceptable. The corresponding RfD and SF values used in equations were shown in Table S4.

### 3 Results and discussion

#### 3.1 Pollution levels and spatial distribution of the heavy metals

Statistical descriptive analysis results of elements including minimum, maximum, mean, standard deviation, and coefficient of variation, in soils from study area are listed in Table 1. The results showed that the arithmetic mean concentrations of the elements were less than the natural background reference values of soils from Henan Province, with the exception of Cd and Pb. However, the peak values for As, Cd, Co, Cu, Pb and Zn were 13.845, 0.744, 11.420, 48.830, 131.860 and 168.760 mg/kg, respectively, and they were 1.2, 10.1, 1.1, 2.5, 6.7 and 2.8 times of the corresponding reference values. Especially, the average values of Cd and Pb were 2.51 and 1.90 times higher than the corresponding background reference values, respectively (Fig.S1), which revealed that heavy metal contamination occurred in the study area. Additionally, Cd levels were above background values in 97.44% of samples, and Pb levels above background values in all samples, demonstrating that Cd and Pb were the dominant concentrated elements in urban soils. The CV shows that the concentration distribution of Cd, Cu, Pb and Zn with considerable high spatial differentiation ( $CV > 30\%$ ), especially Cd with the variation coefficient of 65.05%. The higher variation coefficients may be caused by elevated levels of elements in some sampling sites, suggesting that point sources pollution may exist for Cd, Cu, Pb and Zn in the study area (Huang et al. 2021). The remarkably high kurtosis values of the Pb distribution suggested the occurrence of abnormally high values in some sample, and it was presumably related to anthropogenic discharges (Huang et al. 2020).

In order to clarify the transformation of the heavy metals in the research area, ordinary kriging interpolation was

introduced to obtain the spatial distribution patterns of the target elements. The spatial distribution of As, Cd, Co, Cu, Mn, Ni, Pb, and Zn in downtown area of Zhengzhou were described in Fig. S3. As illustrated in Fig. S3, the spatial distribution maps for As, Co, Mn and Ni were similar, all showing high levels in the northwest and low levels in the southeast. The distribution trends of the four elements were gently and without hot-pot areas (exclude Co). Furthermore, the maximum values of all elements were not significantly different from the background values, and all the mean values were lower than the corresponding background values. The above certificate proves that As, Mn and Ni in urban were uncontaminated, and derive mainly from the soil parent material. A hot-pot area can be observed in contour maps of Co, which locates the intersection of Jinshui Road and Zhongzhou Avenue, Zhengzhou's main thoroughfares, and it is the location of two large commercial plazas (Manhattan Plaza and Chinatown Cultural Plaza) in the area. The slightly elevated Co in this region may be associated with deliberate applications of Co salts, such as Co-containing plating materials in infrastructures, Co treated phosphate fertilizers for greening (Salah et al. 2015; Han et al. 2020). The overall Co distribution map showed that the slight increase of Co concentration in the region has only a minor effect on the mean values, indicating that the Co was also primarily derived from soil parent material.

Spatial distribution patterns of Cd, Cu, Pb and Zn were similar with significant spatial variability as presented in Fig. S3. All hot-pot areas of Cu, Pb and Zn appeared in downtown areas with high traffic volumes or near outer rings of dense traffic lanes, indicating that high concentrations of them might be attributed to engine wear, leaded gasoline, braking and other traffic sources (Dao et al. 2014; Du et al. 2019; Huang et al. 2021). Two hot-pot areas of Cd could be observed in Fig. S3, one near the Zijinshan interchange where traffic volumes were high, and the other

**Table 1** Statistical results of heavy metal concentrations in urban soils from Zhengzhou

	Min. (mg/kg)	Max. (mg/kg)	Mean (mg/kg)	Std. deviation (mg/kg)	Coefficient of variation (%)	Skewness (mg/kg)	Kurtosis (mg/kg)	Background values <sup>a</sup> (mg/kg)
As	6.402	13.845	10.654	1.555	14.60	− 0.215	0.271	11.4
Cd	0.057	0.744	0.186	0.121	65.05	2.261	6.759	0.074
Co	4.340	11.420	7.916	1.191	15.05	0.212	1.060	10.0
Cu	9.350	48.830	15.289	4.993	32.66	4.242	26.081	19.7
Mn	254.160	588.270	374.469	59.281	15.83	1.103	2.385	579
Ni	9.660	24.130	17.583	2.543	14.46	− 0.225	0.424	26.7
Pb	23.450	131.860	37.169	12.454	33.51	5.810	43.510	19.6
Zn	26.880	168.760	52.351	20.745	39.63	2.874	12.265	60.1

<sup>a</sup> Soil background values for heavy metals in Henan Province of China

in the commercial and industrial cluster near Chemical Road, suggesting that Cd pollution was not only related to traffic emissions, but also influenced by commercial activities and industrial discharges (Ma et al. 2016).

### 3.2 Potential ecological risk of heavy metals

The  $I_{geo}$  and  $P_s$  were calculated to evaluation pollution degree of soils in study area, and the results are represented in Fig. 2a and b. As shown in Fig. 2a, the contamination levels of As, Co, Cu, Mn, Ni and Zn were found negligible on the basis of  $I_{geo}$  categories. The average  $I_{geo}$  for Cd and Pb ranged from 0 to 1, and the 75% data points are below 2. Therefore, the levels for Cd and Pb were in the range of uncontaminated and moderately contaminated. As demonstrating in Fig. 2b, the  $P_s$  values ranged from 0.96 to 7.27. According to the classification of  $P_s$  values mentioned before, sample sites in the class of slightly polluted, moderately polluted and seriously polluted domain were 51 (65% of total sites), 15(19% of total sites), 11(14% of total sites), respectively. The results of Nemerow index method indicated that heavy metal pollution in some locations were depressing.

The potential ecological risk indices at different sampling sites assessed according to Eq. (4) was depicted in Fig. 2c. Referring to the  $RI$  category aforementioned, the ecological risks were between low risk ( $RI < 150$ ) and considerable risk ( $300 \leq RI < 600$ ). The  $RI$  of heavy metals at 63 sample points were within 150, demonstrating a low ecological risk for most sample sites. However, 15 sample sites (approximately 20% of total sites) with  $150 \leq RI < 300$  were perceived as moderate risk, and 2

sites with  $RI \geq 300$  were categorized as considerable risk. The Cd contributed the maximum to  $RI$  of all elements, with the exception of sample site 65. The findings of the ecological risk evaluation suggested that contamination control in the study area was imminent, especially for Cd.

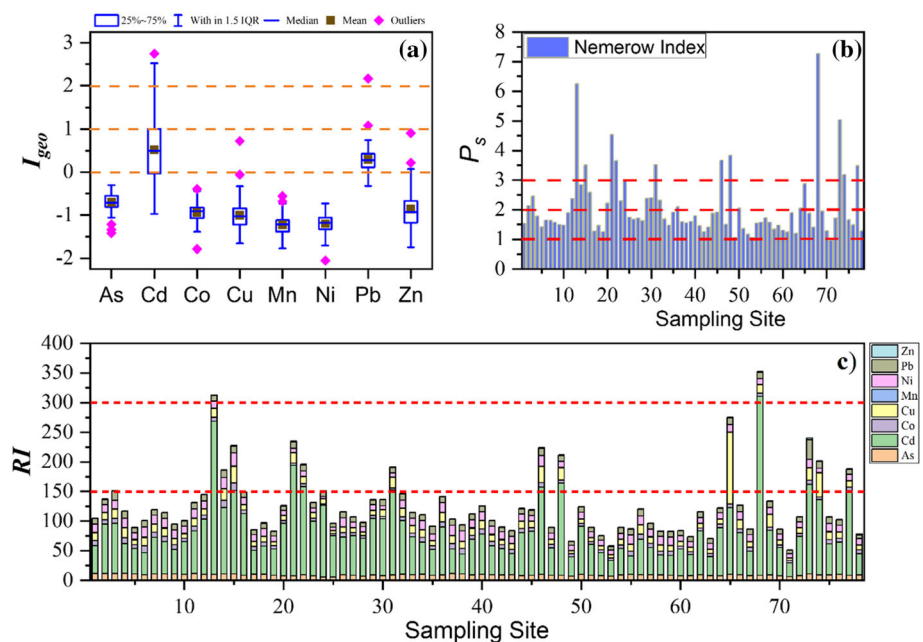
### 3.3 Source allocation and uncertainty assessment

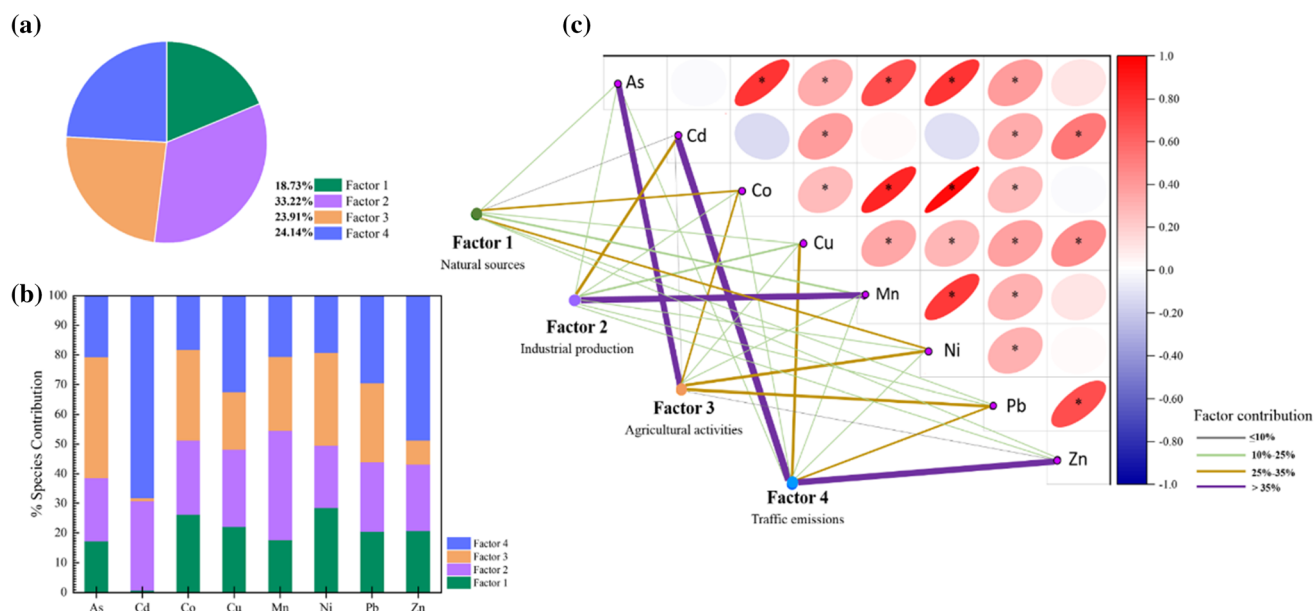
#### 3.3.1 Source allocation of heavy metals

To distinguish the origin and migration of eight elements effectively, Pearson's correlation analysis was performed to ascertain the relationships among the target element concentrations in soils (Fig. 3c) (Ke et al. 2017). Then, the possible sources of contaminants were apportioned and quantified by means of PMF model (Fig. 3a, b). A stable and minimum  $Q$  was introduced in the model to guarantee the residual matrix and the optimum factor number was confirmed as four. The Cd with signal-to-noise ( $S/N$ ) ratios 0.53 was defined as “weak”, and other heavy metals with  $S/N$  ratios above 1 were defined as “strong”. The values of scaled residuals for all heavy metals were between  $-3$  and  $3$ , indicating that there were no abnormal data in the model. The extracted factor loadings and contributions to individual heavy metals by PMF model were presented in Figs. S4 and S5. Further, spatial distributions of elements (Fig. S3) were applied to verify the source identification results.

The results of Pearson's correlation analysis indicated As, Co, Cu, Mn, Ni and Pb were found significantly positively correlated with each other, indicating that they may

**Fig. 2** Pollution levels of heavy metals in urban soil. **a** Box-plots of the geo-accumulation index, **b** Nemerow index and **c** potential ecological risk indices





**Fig. 3** Source apportionment of heavy metals (HMs) in urban soils of the study area. **a** The percentage of contribution for each factor by PMF model. **b** Factor profiles of heavy metals in soils derived from

be from a common origin, while Cd, Cu, Pb and Zn were significantly positively correlated with each other, which may suggest the same source of them (Han et al. 2006). More importantly, the significant correlations of Cu and Pb to other metals indicate multiple sources (Fang et al. 2019). Contribution percentages of the extracted four Factors to the aggregate pollution and individual heavy metals were presented in Fig. 3a and b, respectively.

Factor 1 accounted for 18.73% of total contributions, was characterized by Ni, Co, Cu, Zn, Pb, Mn, As and Cd in descending order (Fig. 3a and b). The Factor profiles were less than 30% for all elements. Furthermore, the observed distribution trend of primarily element Ni was gently without any obvious point source (Fig. S3) and not enriched in any sample site (Fig. S2). Meanwhile, it has been proved to be available in the soil parent materials and pedogenic process in previous studies (Ma et al. 2016; Zhang et al. 2018; Jin et al. 2019; Wang et al. 2019). Therefore, Factor1 may represent the natural source.

Factor 2 was taken 33.22% of total variance, and Mn (36.88%) and Cd (30.01%) received higher weighting than the other elements. Both of the two elements were reported in numerous literatures to be associated with industrial activities, such as smelting, coal consumption, metal processing, galvanization, batteries and plastic application (Li et al. 2017, 2020; Zhang et al. 2018; Men et al. 2020). The proportions of Cd and Mn exceeding background values were 97.44% and 1.28% (Fig. S2), suggesting that the presence of Factor 2 did not have a significant effect on Mn contamination compared to Cd. Two hot-pot areas of Cd

PMF model. **c** Identify the correlations between heavy metals by combining Pearson correlation analysis and PMF model

could be observed in Fig.S3, one near the Zijinshan interchange and the other in the commercial and industrial cluster near Chemical Road, suggesting that Cd pollution was not only related to industrial production, but also influenced by other sources (Ma et al. 2016). Hence, Factor 2 could be interpreted as industrial production.

Factor 3 was weighted primarily on As (40.79%), Ni (31.19%) and Co (30.55%) (Fig. 3b), which made up 23.91% of total contribution (Fig. 3a). Although the mean values of As and Co did not surpass their corresponding local background values, the percentage of sample sites with As and Co exceeding background values were 30.77% and 3.85%, respectively (Fig. S2). Many previous studies confirmed that they were common elements in pesticides and fertilizers (Chen et al. 2016; Zhang et al. 2018; Wang et al. 2019). Pesticides and fertilizers are widely used and overused in urban afforestation and agricultural production (Onakpa et al. 2018). Coincidentally, the spatial distributions of As and Co with high values were in the north region (Fig. S3), which was a typical farming area and had a long history of traditional agricultural of Zhengzhou. As a result, Factor 3 could be assigned to agricultural activities.

Factor 4 consisted predominantly of Cd (68.19%), Zn (48.75%), Cu (32.39%) and Pb (29.49%) (Fig. 3b), was taken 24.14% of total variance (Fig. 3a). Spatial distribution patterns of Cd, Zn, Cu and Pb were similar with significant spatial variability as presented in Fig. S3. Moreover, the proportion of samples that exceed the local background values for Cd, Zn, Cu and Pb were 97.44%,



17.95%, 8.97% and 100%, respectively (Fig. S2). The results showed that Cd and Pb was significantly enriched in urban soil and all of them may be associated with anthropogenic activities. All hot-spot areas of Zn, Cu and Pb could be observed in downtown areas with high traffic volumes or near outer rings of dense traffic lanes, indicating that high concentrations of them might be attributed to leaded petrol, braking, engine wear and other traffic pollution (Du et al. 2019; Huang et al. 2021). Similarly, one hot-spot area of Cd appeared near the Zijinshan interchange where traffic volumes were high, and Cd was described as a component of tyres and lubricants (Huang et al. 2021). As a consequence, Factor 4 was classified as traffic emissions.

### 3.3.2 Uncertainty assessment for factor contributions

In this study, fitted coefficients ( $r^2$ ), variability and *CV* values were applied to test uncertainties of pollutant source analysis. The  $r^2$  between the observed and predicted concentrations were shown in Fig. S6. As illustrated in the figure, the  $r^2$  of all elements were above 0.692, and  $r^2$  values of As, Cd, Co, Cu, Mn, Ni and Zn were even higher than 0.882, which indicated that the PMF model predictions could better explain the values and variation of observed measured values (Men et al. 2019).

The variability of the contribution of Four sources based on BS runs in the PMF model were presented in Fig. S7. Median values in BS runs corresponded to the most accurate Factor contributions, while interquartile ranges represented the variability of Factors' contributions. As shown in Fig. S7, the correlations could be found between median contributions of different slope values and interquartile ranges. The higher  $R^2$ , the lower variability of contributions in bootstrap in different species. The determination coefficient ( $R^2$ ) ranged from 0.984 to 0.997 for all Factors, indicating lower variations of simulated contributions among different heavy metal species.

Fig. S8 showed the base run concentrations of eight heavy metals and *CV* values based on the BS run in every Factor. Generally, higher *CV* values indicate significant deviations. According to the *CV* values in the line chart, Cd exhibited higher *CV* values in Factor 1, Factor 3 and Factor 4, and the *CV* value of As was higher in Factor 2, indicating higher deviation degrees. The *CV* values of Co, Mn and Pb in Factor 1, Cd, Cu, Pb and Zn in Factor 2, Pb and Ni in Factor 3, Cu, Pb, Co and Ni in Factor 4 were all less than 30%, indicating lower deviation degrees. Moreover, higher *CV* values tended to correlate with lower concentration in the base run of the Four Factors, which might be related to the re-sampling process based on BS simulation, or to its own low determined values (Brown et al. 2015). Despite

the fact that the BS results could not reflect the uncertainty of rotation normally, they can identify the factors with poorly reproducible (Xie et al. 2013). In addition, the lower concentrations of heavy metal were allocated by the PMF model, the less impact on the results of the PMF base run (Liu et al. 2018). Therefore, it could be concluded that all four source Factors allocated by the PMF model were reliable.

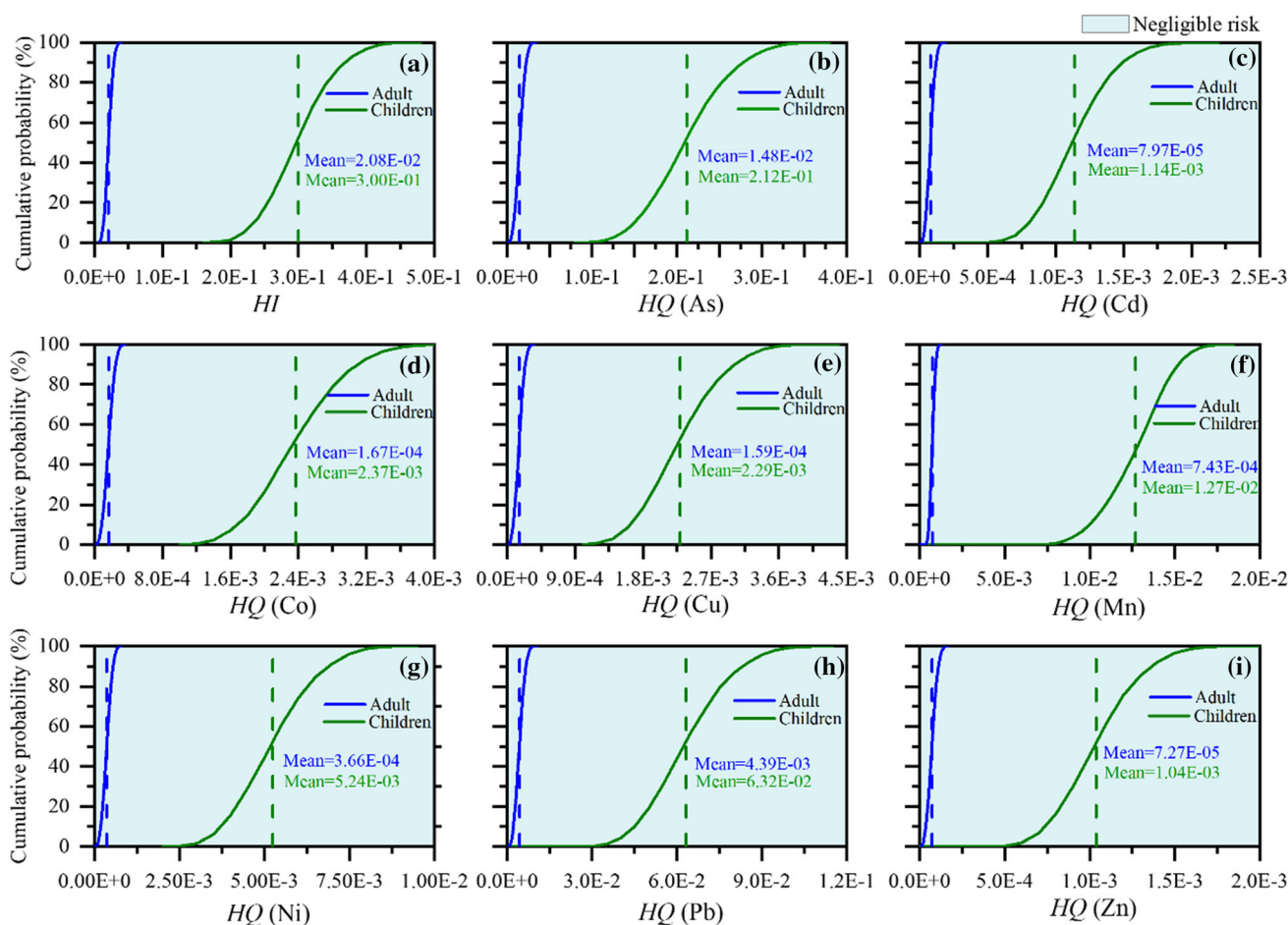
## 3.4 Probabilistic health risk assessment

### 3.4.1 Concentration-oriented health risk analysis

In order to further clarify the health effects of pollutants on the local residents, the non-carcinogenic and carcinogenic risk base on heavy metal concentrations were evaluated, and the results were demonstrated in the Figs. 4 and 5.

As presented in Fig. 4, the *HI* and *HQ* of individual element for both children and adult were no more than the USEPA's specified value of 1, suggesting that the potential non-carcinogenic risks could be ignored. The mean *HQ* values of heavy metals exhibited a trend of As > Pb > Mn > Ni > Co  $\approx$  Cu > Cd > Zn. Among three exposure routes, the ingestion way was the primary contributor of potential risk (Table S5). Additionally, compared with adults, children were more susceptible to non-carcinogenic risks. Coincidentally, this observation was also previously reported in studies (Tong et al. 2019; Huang et al. 2021), which alerted the public to the benefits of personal hygiene, reducing the frequency of hand-mouth contact and taking appropriate skin protection measures outdoors, especially for developing children.

The probability distribution of *TCR* in Fig. 5 revealed that the carcinogenic risk of five elements in the studied area cannot be negligible. As showed in Fig. 5a, the mean values of *TCR* caused by heavy metal pollutants for adults ( $1.03E-04$ ) and children ( $1.28E-04$ ) all exceed the upper acceptable limit of carcinogenic risk ( $1E-04$ ). Further, the high exceedance percentage of threshold were found in Fig. S9, indicating that both adults and children were more likely to cause cancer and pose an unacceptable risk. The *CR* values of heavy metals were different from each other, and they could be available in decreasing order of Ni, Co, As, Cd and Pb (Fig. 5b–f). The proportion of risk values for Ni, Co and As above the threshold of  $1E-04$  could be found in Fig. S9. Besides, the mean *CR* values of Ni, Co and As were all above the acceptable limit of  $1E-06$  to the extent that they were at the 5th percentile (Table S6), and they were considered to be principal contributor to cancer risk. Therefore, those elements were considered as priority pollutants for further risk control.



**Fig. 4** Probability distribution of non-carcinogenic hazard index for **a** *HI* and *HQ* of **b** As, **c** Cd, **d** Co, **e** Cu, **f** Mn, **g** Ni, **h** Pb and **i** Zn, the vertical dashed lines with blue or green colour represented the mean values for adults and children, respectively

### 3.4.2 Source-oriented health risk analysis

To differentiate the contribution of different heavy metal pollution sources to health risk, a comprehensive method on the basis of probabilistic health risk assessment developed by Huang (Huang et al. 2021) was adopted to identify source-oriented health risk. As depicted in Fig. 6, the risk profiles of the four sources showed similar distributional characteristics between children and adults, and agricultural pollution was considered to be the dominant source of health risks for the two groups. According to Fig. 4, As was identified as the dominant element for non-carcinogenic health risks owing to its highest mean *HQ* value. The Fig. 5 showed that Ni, Co and As were the main elements causing carcinogenic risks. The three elements mentioned above held relatively high weightings in Factor 3 (Fig. S5), and they were distributed highly in the north of the city (Fig. S3). In consequence, reducing emissions of heavy metals from agricultural activities is an effective step in decreasing potential risk to the public.

## 4 Conclusions

The contaminated level, source recognition and risk evaluation of eight elements in urban soils from Zhengzhou city were carried out in the study. Results found that Cd and Pb were the dominant concentrated contaminants and presented higher ecological risks compared to the others. The probability distribution of concentration-oriented health risk revealed that adverse health effect of heavy metals for all populations was unlikely, while the carcinogenic risk produced by five elements could not be acceptable. Moreover, the children were more susceptible to health risks than adults, and the ingestion way was considered to be the major route of risk exposure. Four Factors were identified based on the conjunction of PMF model and correlation analysis. Nevertheless, sampling locations in this study were mostly close to traffic lanes, which may interfere with the source analysis results to some extent. Uncertainty analysis revealed that there were certain uncertainties in PMF source analysis, but pollution

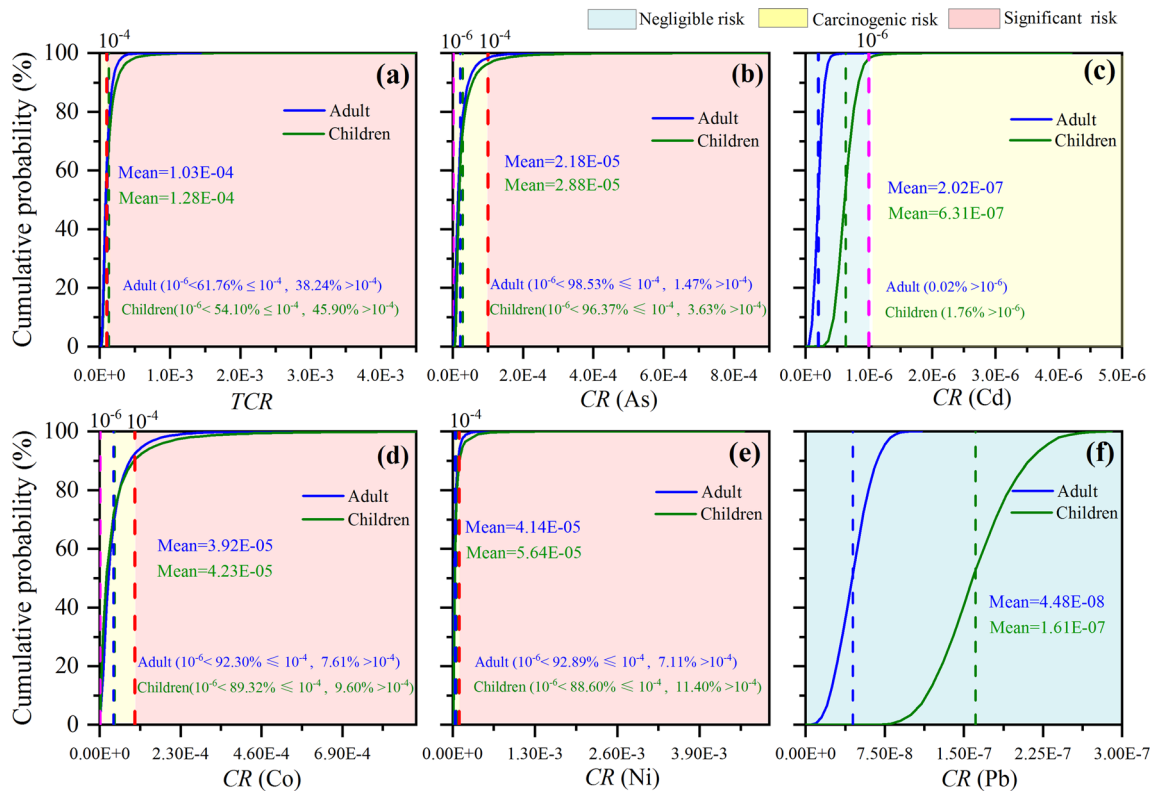
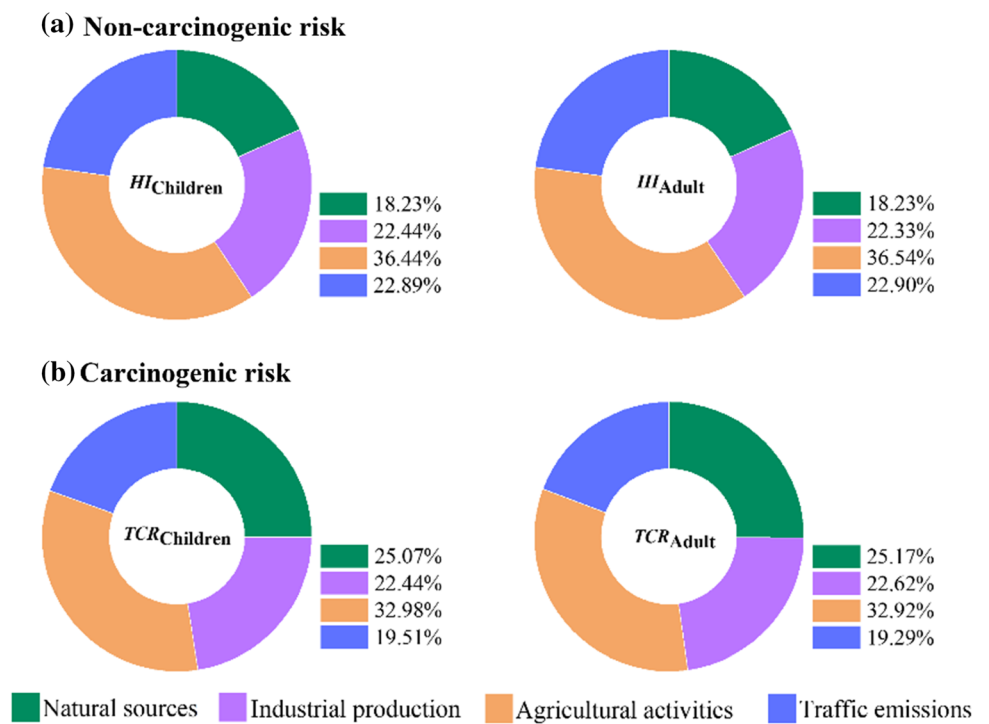


Fig. 5 Probability distribution and the percentage surpassed 1E-06 for a TCR and CR of b As, c Cd, d Co, e Ni and f Pb

Fig. 6 Comparison of health risks caused by different pollution sources of heavy metals. a Non-carcinogenic risk and b Carcinogenic risk for children, adult, respectively



source Factors still retained reliable. Subsequently, the source-oriented health risk was carried out through PMF model in conjunction with Monte Carlo simulation technique, and the agricultural source was confirmed to be the main source of health risks, and As, Co and Ni were identified as the highlight risky elements, which should be given special consideration for further risk control.

**Author contributions** All authors contributed to the study conception and design. Experimental design and material collection were conducted by JZ, EG and YC. The final version of all data reviewed and approved by LL, CN and YY. The research supervision and manuscript writing were performed by LM. All authors read and approved the final manuscript.

**Funding** This work was supported by Key Laboratory for Monitor and Remediation of Heavy Metal Polluted Soils of Henan Province (2017207) and Henan Provincial Science and Technology Department (222102320396).

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Ethical approval** All ethical responsibilities are considered regarding the publication of this paper.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s00477-022-02297-w>.

## References

- Awasthi AK, Zeng XL, Li JH (2016) Environmental pollution of electronic waste recycling in India: a critical review. *Environ Pollut* 211:259–270. <https://doi.org/10.1016/j.envpol.2015.11.027>
- Brown SG, Eberly S, Paatero P, Norris GA (2015) Methods for estimating uncertainty in PMF solutions: examples with ambient air and water quality data and guidance on reporting PMF results. *Sci Total Environ* 518–519:626–635. <https://doi.org/10.1016/j.scitotenv.2015.01.022>
- Chai L, Wang Y, Wang X, Ma L, Cheng Z, Su L, Liu M (2021) Quantitative source apportionment of heavy metals in cultivated soil and associated model uncertainty. *Ecotoxicol Environ Saf* 215:112150. <https://doi.org/10.1016/j.ecoenv.2021.112150>
- Chen T, Chang Q, Liu J, Clevers J, Kooistra L (2016) Identification of soil heavy metal sources and improvement in spatial mapping based on soil spectral information: a case study in northwest China. *Sci Total Environ* 565:155–164. <https://doi.org/10.1016/j.scitotenv.2016.04.163>
- Chen RH, Chen HY, Song LT, Yao ZP, Meng FS, Teng YG (2019) Characterization and source apportionment of heavy metals in the sediments of Lake Tai (China) and its surrounding soils. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2019.133819>
- Cui X, Geng Y, Sun R, Xie M, Feng X, Li X, Cui Z (2021) Distribution, speciation and ecological risk assessment of heavy metals in Jinan Iron & Steel Group soils from China. *J Clean Prod*. <https://doi.org/10.1016/j.jclepro.2021.126504>
- Dao L, Morrison L, Zhang H, Zhang C (2014) Influences of traffic on Pb, Cu and Zn concentrations in roadside soils of an urban park in Dublin, Ireland. *Environ Geochem Health* 36:333–343. <https://doi.org/10.1007/s10653-013-9553-8>
- Dehghan Madiseh S, Savary A, Parham H, Sabzalizadeh S (2009) Determination of the level of contamination in Khuzestan coastal waters (Northern Persian Gulf) by using an ecological risk index. *Environ Monit Assess* 159(1–4):521–530. <https://doi.org/10.1007/s10661-008-0648-0>
- Du X, Zhu Y, Han Q, Yu Z (2019) The influence of traffic density on heavy metals distribution in urban road runoff in Beijing China. *Environ Sci Pollut Res Int* 26(1):886–895. <https://doi.org/10.1007/s11356-018-3685-4>
- Faisal M, Wu ZN, Wang HL, Hussain Z, Shen CY (2021) Geochemical mapping, risk assessment, and source identification of heavy metals in road dust using positive matrix factorization (PMF). *Atmosphere-Basel*. <https://doi.org/10.3390/atmos12050614>
- Fang XH, Peng B, Wang X, Song ZL, Zhou DX, Wang Q, Qin ZL, Tan CY (2019) Distribution, contamination and source identification of heavy metals in bed sediments from the lower reaches of the Xiangjiang River in Hunan province, China. *Sci Total Environ* 689:557–570. <https://doi.org/10.1016/j.scitotenv.2019.06.330>
- Gu YG, Gao YP (2018) Bioaccessibilities and health implications of heavy metals in exposed-lawn soils from 28 urban parks in the megacity Guangzhou inferred from an in vitro physiologically-based extraction test. *Ecotoxicol Environ Saf* 148:747–753. <https://doi.org/10.1016/j.ecoenv.2017.11.039>
- Guan QY, Wang FF, Xu CQ, Pan NH, Lin JK, Zhao R, Yang YY, Luo HP (2018) Source apportionment of heavy metals in agricultural soil based on PMF: a case study in Hexi Corridor, northwest China. *Chemosphere* 193:189–197. <https://doi.org/10.1016/j.chemosphere.2017.10.151>
- Hakanson L (1980) An ecological risk index for aquatic pollution control. a sedimentological approach. *Water Res* 14(8):975–1001
- Han YM, Du PX, Cao JJ, Posmentier ES (2006) Multivariate analysis of heavy metal contamination in urban dusts of Xi'an. *Central China Sci Total Environ* 355(1–3):176–186. <https://doi.org/10.1016/j.scitotenv.2005.02.026>
- Han Q, Wang M, Cao J, Gui C, Liu Y, He X, He Y, Liu Y (2020) Health risk assessment and bioaccessibilities of heavy metals for children in soil and dust from urban parks and schools of Jiaozuo China. *Ecotoxicol Environ Saf* 191:110157. <https://doi.org/10.1016/j.ecoenv.2019.110157>
- Hou SN, Zheng N, Tang L, Ji XF, Li YY, Hua XY (2019) Pollution characteristics, sources, and health risk assessment of human exposure to Cu, Zn, Cd and Pb pollution in urban street dust across China between 2009 and 2018. *Environ Int* 128:430–437. <https://doi.org/10.1016/j.envint.2019.04.046>
- Huang YN, Dang F, Li M, Zhou DM, Song Y, Wang JB (2020) Environmental and human health risks from metal exposures nearby a Pb-Zn-Ag mine, China. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2019.134326>
- Huang J, Wu Y, Sun J, Li X, Geng X, Zhao M, Sun T, Fan Z (2021) 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* 415:125629. <https://doi.org/10.1016/j.jhazmat.2021.125629>

- Islam ARMT, Islam HMT, Mia MU, Khan R, Habib MA, Bodrud-Doza M, Siddique MAB, Chu R (2020) Co-distribution, possible origins, status and potential health risk of trace elements in surface water sources from six major river basins, Bangladesh. *Chemosphere* 249:126180. <https://doi.org/10.1016/j.chemosphere.2020.126180>
- Jaishankar M, Tseten T, Anbalagan N, Mathew BB, Beeregowda KN (2014) Toxicity, mechanism and health effects of some heavy metals. *Interdiscip Toxicol* 7(2):60–72. <https://doi.org/10.2478/intox-2014-0009>
- Jiang CL, Zhao Q, Zheng LG, Chen X, Li C, Ren MX (2021) Distribution, source and health risk assessment based on the Monte Carlo method of heavy metals in shallow groundwater in an area affected by mining activities China. *Ecotox Environ Safe*. <https://doi.org/10.1016/j.ecoenv.2021.112679>
- Jin Y, O'Connor D, Ok YS, Tsang DCW, Liu A, Hou D (2019) Assessment of sources of heavy metals in soil and dust at children's playgrounds in Beijing using GIS and multivariate statistical analysis. *Environ Int* 124:320–328. <https://doi.org/10.1016/j.envint.2019.01.024>
- Ke X, Gui SF, Huang H, Zhang HJ, Wang CY, Guo W (2017) Ecological risk assessment and source identification for heavy metals in surface sediment from the Liaohe River protected area, China. *Chemosphere* 175:473–481. <https://doi.org/10.1016/j.chemosphere.2017.02.029>
- Lee MY, Kim YJ, Hwang DG, Kang YY, Shin SK, Jeon TW (2021) Potential risk of exposure to heavy metals from co-processing of secondary wastes in the Republic of Korea. *J Environ Manage* 286:112164. <https://doi.org/10.1016/j.jenvman.2021.112164>
- Li Z, Ma Z, van der Kuijp TJ, Yuan Z, Huang L (2014) A review of soil heavy metal pollution from mines in China: pollution and health risk assessment. *Sci Total Environ* 468–469:843–853. <https://doi.org/10.1016/j.scitotenv.2013.08.090>
- Li HH, Chen LJ, Yu L, Guo ZB, Shan CQ, Lin JQ, Gu YG, Yang ZB, Yang YX, Shao JR, Zhu XM, Cheng Z (2017) Pollution characteristics and risk assessment of human exposure to oral bioaccessibility of heavy metals via urban street dusts from different functional areas in Chengdu, China. *Sci Total Environ* 586:1076–1084. <https://doi.org/10.1016/j.scitotenv.2017.02.092>
- Li Y, Yuan Y, Sun C, Sun T, Liu X, Li J, Fang L, Fan Z (2020) Heavy metals in soil of an urban industrial zone in a metropolis: risk assessment and source apportionment. *Stoch Env Res Risk A* 34(2):435–446. <https://doi.org/10.1007/s00477-020-01779-z>
- Liu Q, Cao J, Li KQ, Miao XH, Li G, Fan FY, Zhao YC (2009) Chromosomal aberrations and DNA damage in human populations exposed to the processing of electronics waste. *Environ Sci Pollut R* 16:329–338. <https://doi.org/10.1007/s11356-008-0087-z>
- Liu R, Men C, Yu W, Xu F, Wang Q, Shen Z (2018) Uncertainty in positive matrix factorization solutions for PAHs in surface sediments of the Yangtze River Estuary in different seasons. *Chemosphere* 191:922–936. <https://doi.org/10.1016/j.chemosphere.2017.10.070>
- Liu S, Pan G, Zhang Y, Xu J, Ma R, Shen Z, Dong S (2019a) Risk assessment of soil heavy metals associated with land use variations in the riparian zones of a typical urban river gradient. *Ecotoxicol Environ Saf* 181:435–444. <https://doi.org/10.1016/j.ecoenv.2019.04.060>
- Liu Y, Zhang Y, Li J, Wu N, Li W, Niu Z (2019b) Distribution, partitioning behavior and positive matrix factorization-based source analysis of legacy and emerging polyfluorinated alkyl substances in the dissolved phase, surface sediment and suspended particulate matter around coastal areas of Bohai Bay, China. *Environ Pollut* 246:34–44. <https://doi.org/10.1016/j.envpol.2018.11.113>
- Ma L, Yang ZG, Li L, Wang L (2016) Source identification and risk assessment of heavy metal contaminations in urban soils of Changsha, a mine-impacted city in Southern China. *Environ Sci Pollut R* 23(17):17058–17066. <https://doi.org/10.1007/s11356-016-6890-z>
- Ma L, Yang ZG, Kong Q, Wang L (2017) Extraction and determination of arsenic species in leafy vegetables: Method development and application. *Food Chem* 217:524–530. <https://doi.org/10.1016/j.foodchem.2016.09.015>
- Men C, Liu R, Xu F, Wang Q, Guo L, Shen Z (2018) Pollution characteristics, risk assessment, and source apportionment of heavy metals in road dust in Beijing, China. *Sci Total Environ* 612:138–147. <https://doi.org/10.1016/j.scitotenv.2017.08.123>
- Men C, Liu R, Wang Q, Guo L, Miao Y, Shen Z (2019) Uncertainty analysis in source apportionment of heavy metals in road dust based on positive matrix factorization model and geographic information system. *Sci Total Environ* 652:27–39. <https://doi.org/10.1016/j.scitotenv.2018.10.212>
- Men C, Liu RM, Xu LB, Wang QR, Guo LJ, Miao YX, Shen ZY (2020) Source-specific ecological risk analysis and critical source identification of heavy metals in road dust in Beijing China. *J Hazard Mater*. <https://doi.org/10.1016/j.jhazmat.2019.121763>
- Mihankhah T, Saeedi M, Karbassi A (2020) A comparative study of elemental pollution and health risk assessment in urban dust of different land-uses in Tehran's urban area. *Chemosphere*. <https://doi.org/10.1016/j.chemosphere.2019.124984>
- Onakpa MM, Njan AA, Kalu OC (2018) A review of heavy metal contamination of food crops in Nigeria. *Ann Glob Health* 84(3):488–494
- Paatero P, Eberly S, Brown SG, Norris GA (2014) Methods for estimating uncertainty in factor analytic solutions. *Atmosph Measur Tech* 7(3):781–797. <https://doi.org/10.5194/amt-7-781-2014>
- Pecina V, Brtnicky M, Baltazar T, Juricka D, Kynicky J, Vasinova Galiova M (2021) Human health and ecological risk assessment of trace elements in urban soils of 101 cities in China: a meta-analysis. *Chemosphere* 267:129215. <https://doi.org/10.1016/j.chemosphere.2020.129215>
- Qiao J, Zhu Y, Jia X, Shao M, Niu X, Liu J (2020) Distributions of arsenic and other heavy metals, and health risk assessments for groundwater in the Guanzhong Plain region of China. *Environ Res* 181:108957. <https://doi.org/10.1016/j.envres.2019.108957>
- Safur Rahman M, Khan MDH, Jolly YN, Kabir J, Akter S, Salam A (2019) Assessing risk to human health for heavy metal contamination through street dust in the Southeast Asian Megacity: Dhaka, Bangladesh. *Sci Total Environ* 660:1610–1622. <https://doi.org/10.1016/j.scitotenv.2018.12.425>
- Salah EA, Turki AM, Mahal SN (2015) Chemometric evaluation of the heavy metals in urban soil of Fallujah City Iraq. *J Environ Protect* 06(11):1279–1292. <https://doi.org/10.4236/jep.2015.61112>
- Siddiqui AU, Jain MK, Masto RE (2020) Pollution evaluation, spatial distribution, and source apportionment of trace metals around coal mines soil: the case study of eastern India. *Environ Sci Pollut Res Int* 27(10):10822–10834. <https://doi.org/10.1007/s11356-019-06915-z>
- Siddiqui AU, Jain MK, Masto RE (2021) Distribution of some potentially toxic elements in the soils of the Jharia Coalfield: a probabilistic approach for source identification and risk assessment. *Land Degrad Dev* 33(2):333–345. <https://doi.org/10.1002/ldr.4155>
- Srivastava D, Xu JS, Vu TV, Liu D, Li LJ, Fu PQ, Hou SQ, Palmerola NM, Shi ZB, Harrison RM (2021) Insight into PM2.5 sources by applying positive matrix factorization (PMF) at urban and rural sites of Beijing. *Atmos Chem Phys* 21(19):14703–14724
- Sun L, Guo D, Liu K, Meng H, Zheng Y, Yuan F, Zhu G (2019) Levels, sources, and spatial distribution of heavy metals in soils

- from a typical coal industrial city of Tangshan, China. *CATENA* 175:101–109. <https://doi.org/10.1016/j.catena.2018.12.014>
- Tong R, Jia Q, Ma X, Fang Y, Wang W (2019) Comprehensive comparison of probabilistic health risks of soil heavy metals in China's mining areas. *Hum Ecol Risk Assess* 26(8):2059–2077. <https://doi.org/10.1080/10807039.2019.1649115>
- Tong S, Li H, Wang L, Tudi M, Yang L (2020) Concentration, spatial distribution, contamination degree and human health risk assessment of heavy metals in urban soils across China between 2003 and 2019-A systematic review. *Int J Environ Res Public Health*. <https://doi.org/10.3390/ijerph17093099>
- USEPA (2014) EPA Positive matrix factorization (PMF) 5.0 fundamentals and user guide. U.S. Environment Protection Agency (Washington DC)
- Varela JP, Valente AJM, Duraes L (2019) Assessment of heavy metal pollution from anthropogenic activities and remediation strategies: a review. *J Environ Manage* 246:101–118. <https://doi.org/10.1016/j.jenvman.2019.05.126>
- Wang G, Liu HQ, Gong Y, Wei Y, Miao AJ, Yang LY, Zhong H (2017) Risk assessment of metals in urban soils from a typical industrial city Suzhou Eastern China. *Int J Environ Res Public Health*. <https://doi.org/10.3390/ijerph14091025>
- Wang S, Cai LM, Wen HH, Luo J, Wang QS, Liu X (2019) Spatial distribution and source apportionment of heavy metals in soil from a typical county-level city of Guangdong Province, China. *Sci Total Environ* 655:92–101. <https://doi.org/10.1016/j.scitotenv.2018.11.244>
- Wang H, Shen C, Kang Y, Deng Q, Lin X (2020a) Spatial distribution of pollution characteristics and human health risk assessment of exposure to heavy elements in road dust from different functional areas of Zhengzhou China. *Environ Sci Pollut Res Int* 27(21):26650–26667. <https://doi.org/10.1007/s11356-020-08942-7>
- Wang J, Shan Q, Liang X, Guan F, Zhang Z, Huang H, Fang H (2020b) Levels and human health risk assessments of heavy metals in fish tissue obtained from the agricultural heritage rice-fish-farming system in China. *J Hazard Mater* 386:121627. <https://doi.org/10.1016/j.jhazmat.2019.121627>
- Wu W, Wu P, Yang F, Sun DL, Zhang DX, Zhou YK (2018) Assessment of heavy metal pollution and human health risks in urban soils around an electronics manufacturing facility. *Sci Total Environ* 630:53–61. <https://doi.org/10.1016/j.scitotenv.2018.02.183>
- Xie M, Piedrahita R, Dutton SJ, Milford JB, Hemann JG, Peel JL, Miller SL, Kim SY, Vedal S, Sheppard L, Hannigan MP (2013) Positive matrix factorization of a 32-month series of daily PM<sub>2.5</sub> speciation data with incorporation of temperature stratification. *Atmos Environ*. <https://doi.org/10.1016/j.atmosenv.2012.09.034>
- Zeng S, Ma J, Yang Y, Zhang S, Liu GJ, Chen F (2019) Spatial assessment of farmland soil pollution and its potential human health risks in China. *Sci Total Environ* 687:642–653. <https://doi.org/10.1016/j.scitotenv.2019.05.291>
- Zhang XW, Wei S, Sun QQ, Wadood SA, Guo BL (2018) Source identification and spatial distribution of arsenic and heavy metals in agricultural soil around Hunan industrial estate by positive matrix factorization model, principle components analysis and geo statistical analysis. *Ecotox Environ Safe* 159:354–362. <https://doi.org/10.1016/j.ecoenv.2018.04.072>
- Zhang ZY, Mamat A, Simayi Z (2019) Pollution assessment and health risks evaluation of (metalloid) heavy metals in urban street dust of 58 cities in China. *Environ Sci Pollut R* 26(1):126–140. <https://doi.org/10.1007/s11356-018-3555-0>
- Zhang H, Cheng SQ, Li HF, Fu K, Xu Y (2020) Groundwater pollution source identification and apportionment using PMF and PCA-APCA-MLR receptor models in a typical mixed land-use area in Southwestern China. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2020.140383>
- Zhang H, Zhan Y, Li J, Chao CY, Liu Q, Wang C, Jia S, Ma L, Biswas P (2021) Using Kriging incorporated with wind direction to investigate ground-level PM<sub>2.5</sub> concentration. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2020.141813>

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.