Analysis of Heavy Metal Pollution in Urban Topsoil

Shaojie Kang[1, a], Lianzheng Chen[2,b] and Yang Qiu[3,c]

[1]College of Computer Science & Technology,Wuhan University of Technology,Peace Avenue 1040, Wuhan, Hubei Province, China

[2]School of Logistics Engineering,Wuhan University of Technology,Peace Avenue 1040, Wuhan, Hubei Province, China

[3]School of Transportation,Wuhan University of Technology,Peace Avenue 1040, Wuhan, Hubei Province, China

[a]283381851@qq.com, [b]1002830149@qq.com, [c] 657852355@qq.com

Keywords: Fuzzy recognition; Concentration-displacement offset model; Cubic interpolation; Propagation.

Abstract. Evaluation of geological environment in urban soil is an effective way to indicate the quality of the urban environment. In a city, after an acquisition of sufficient samples of topsoil concentrations of heavy metals information, we can conduct a numerical analysis of the city comprehensively and explore the locations of pollution source of each heavy metal. Then we apply the fuzzy recognition method to plot eight kinds of spatial distribution of heavy metal elements in the urban areas and their degrees of pollution. According to the national environmental quality standards for soil, the method can be used to determine the pollution levels for each sample point and different levels of pollution in life zones, industrial parks, mountains, trunk roads and park green areas. Through the analysis of the propagation features of heavy metal pollutants, we can establish a concentration-displacement offset model to determine the location of each source.

Introduction

In accordance with the functional division, a city can generally be divided into life zones, industrial parks, mountains, trunk roads and park green areas.Different regional environments under the effects of human activities varies to different degrees.

Firstly, we divide the city into lattice spacing about 1 km grid area, in accordance with 1 sampling point on surface soil per square kilometer (0~10 cm depth), and number a sample point, and use the GPS to record the location of it. With the help of specialized instrumental test and analysis, concentration of several elements in each sample will be easily acquired. What's more, by taking samples per 2 kms of the areas that are remote and the industry areas, we will get the background value of elements in the topsoil in the area[1].

What we need to do is to attain the spatial distribution of eight major heavy metal elements in the city, and analyzing their pollution levels in different areas respectively. Finally, analysis of propagation features of heavy metal pollutants will be made, which is the foundation of an effective model to determine the location of sources of pollution.

Locations of Pollution Sources

In order to make the pollution degree more standard, we use the optimal weight method to establish fuzzy recognition model to define the similarity degree of pollution degree between each sample point and the standard soil, which will help to analyze the pollution degree of different areas.

In order to eliminate the effects of the difference of magnitude of indicators, and unify the changing range of each indicator [3], there's a need to perform standardization. Standardized matrix of topsoil is $X = (x_{ij})_{n \times m}$.

Then convert X into fuzzy evaluation matrix,

$$R = (r_{ij})_{n\times m} \quad . \tag{1}$$

The smaller the contamination indicator is, the better environment the area has, therefore it yields,

$$r_{ij} = \frac{x_{\max}(j) - x_{ij}}{x_{\max}(j) - x_{\min}(j)} \cdot \tag{2}$$

And $x_{\max}(j)$ and $x_{\min}(j)$ are the maximum and minimum values of each sample point in the index j respectively.

We can use the optimal weight method on the basis of a weighted linear function of the dimensionless index, and attempt to maximize the sample variance of the function reaches its maximum by optimizing the weight coefficient. The basic idea is to seek the optimal solution of the weight vector Eq. 3, so that the sample variance can reach its maximum under certain constraint.

$$\max \quad s^2 = w^T V w \quad st. \quad w^T w = 1 \quad . \tag{3}$$

Where, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2$, $\bar{u} = \frac{1}{m}\sum_{j=1}^{m} u_j$. And u_j is the similarity degree of the sample points with standard soil class j, V is the covariance matrix of standardized matrix R,

$$V = (V_{ij})_{n\times n} \quad . \tag{4}$$

In the Eq.4, $v_{ij} = \frac{1}{n}\sum_{k=1}^{m}(r_{ki} - \bar{r}_i)(r_{kj} - \bar{r}_j),\ i, j = 1, 2, \cdots, m$, $\bar{r}_i = \frac{1}{m}\sum_{i=1}^{m} r_{ij}$, $\bar{r}_j = \frac{1}{m}\sum_{j=1}^{m} r_{ij}$.

In fact, the optimal solution of weighting problem in Eq.4 is unit feature root vectors of the maximum feature root of the sample covariance matrix: $w = (w_1, w_2, w_3, \cdots, w_n)^T$.

Then we can calculate the soil quality grade of each sampling point. Assuming that the indicator of k-th sample point is e_k : $e_k = (e_{k1}, e_{k2}, \cdots, e_{kn})^T$ .It can be standardized into $f : f_k = (f_{k1}, f_{k2}, \cdots, f_{kn})^T$. Euclidean distance (p=2) is used to calculate the similarity degree,

$$u_{kh} = \left| \sum_{j=1}^{m} \frac{\sum_{i=1}^{n}(w_i(f_{ki} - r_{ih}))^2}{\sum_{i=1}^{n}(w_i(f_{ii} - r_{ij}))^2} \right|^{-1} , h = 1, 2, \cdots, m . \tag{5}$$

The similarity degree vector can be calculated for each sample point and all kinds of standard soil by using Eq.5,

$$u_k = (u_{k1}, u_{k2}, \cdots, u_{km})^T, k = 1, 2, \cdots, p \quad . \tag{6}$$

The contamination type of each sample point can be determined by the similarity degree vector and the closest principle. Through the test of all kinds of soil heavy metal, the concentration and background values of arsenic, cadmium, chromium, copper, mercury, nickel, lead and zincates in the sampling points can be calculated. The eight kinds of metal will be selected as the evaluation indicators of soil environmental quality.

The standardized matrix R which is the soil pollution grade of assessment can be calculated by using Eq.2 and Eq.3.It's as follows,

$$R = \begin{pmatrix} 1.000 & 0.859 & 1.000 & 1.000 \\ 1.053 & 1.008 & 1.110 & 1.085 \\ 0.581 & 0.101 & 0.370 & 0.435 \\ 0.853 & 0.657 & 0.741 & 0.217 \\ 1.055 & 1.009 & 1.109 & 1.084 \\ 0.867 & 0.606 & 0.926 & 0.652 \\ 0.581 & 0.657 & 0.000 & 0.000 \\ 0.000 & 0.000 & 0.185 & 0.000 \end{pmatrix}$$

According to Eq.4, it's easy to get covariance matrix V . The largest characteristic root of matrix V and the corresponding normalization of the eigenvectors are as follows:

$$\lambda = 0.1698 \qquad w = (-0.078, -0.105, -0.239, -0.128, -0.137, -0.10, -0.10, -0.110)^T$$

Then we can get the similarity degree vector corresponding to all types of standard soils. And the contamination type of each sample point can be determined by the closest principle.The statistics result is in the following Table 1.From this table, we can conclude that the degree of contamination of the areas listed in an ascending order is: mountains<green park<life zones <trunk roads <industrial areas.

Table 1. Statistics Result in Each Area

	Cleanness		Minor contamination		Moderate contamination		Severe contamination		Percent of contamination points
	Sum	Percent	Sum	Percent	Sum	Percent	Sum	Percent	
Life zones	25	56.82%	15	34.09%	2	4.55%	2	4.55%	43.18%
Industrial areas	12	33.33%	18	50.00%	3	8.33%	3	8.33%	66.67%
Mountains	58	87.88%	8	12.12%	0	0.00%	0	0.00%	12.12%
Trunk roads	65	47.10%	52	37.68%	14	10.14%	7	5.07%	52.90%
Green park	27	77.14%	6	17.14%	1	2.86%	1	2.86%	22.86%

The Spatial Distribution of Heavy Metal

In this chapter we will introduce the principle of the interpolation method[4].In this function $z = f(x, y)$,the height of the sampling point is determined by plane coordinate x, y . The function value in 319 sampling points $(x_0, y_0), \cdots, (x_{318}, y_{318})$ is given in the data in the attachment. And x_k ranges from 0 to 28654 and y_k ranges from 0 to 18449.Furthermore,

$$0 \le x_0 < x_1 < \cdots < x_{318} \le 28654,\ 0 \le y_0 < y_1 < \cdots < y_{318} \le 18449,\ z_k = f(x_k, y_k)$$

Then a 318 order polynomial $P(x, y)$ over 319 points can be constructed. In the polynomial, there's only a need to know the value of x_k, y_k and z_k , without a high order numerical derivative. In the interval [0,28654], [0,18449] of x, y , the value $f(x, y)$ can be approximated with polynomial $P(x, y)$ .However, if the error function is required,it is necessary to know the range of the value $f^{319}(x, y): M = \max\left\{\left|f^{319}(x, y)\right| : 0 \le x \le 28654, 0 \le y \le 18449\right\}$ .The spatial distribution is as follows,

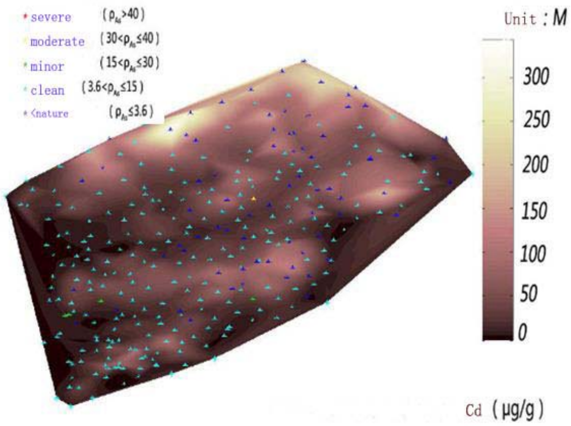

Fig.1 Spatial Concentration Distribution of Cd Metal

The Propagation Features of Heavy Metal Pollutants

Definition: the concentration - displacement offset refers to the concept that in the three-dimensional Cartesian coordinate system, there's an existence of a variable (x、y、z) as the dependent variable, the other two variables as a function of the dependent variable, in which the variable is only with the other two variables change. To this problem, in the case of function $z = f(x, y)$, z_1 and z_2 are respectively some kind of metal element's concentration at the corresponding point in surface spatial coordinates (x, y) [4]. Therefore, the concentration - displacement offset is $\Delta(x, y) = z_1 - z_2$.

It is obtained that at each sample point location, the concentration of the metal elements can be expressed as $\rho = z_1 - \Delta$.

By using the interpolation method, which has obtained the surface distribution of the region, the distribution of concentration- displacement offset in this region will be easily obtained. In order to achieve the concentration distribution and facilitate subsequent calculations, we make full use of the discrete data analyzing approach, and establish the concentration-displacement offset model.Detailed operations are as follows:

- After trial and error, take 10 meters in the latitude and longitude as the step to to divide the region into grids.That's $x_i = 1:10:\max(x)$, $y_i = 1:10:\max(y)$.
- For these eight kinds of metal elements, respectively, make the concentration - displacement offset matrix z_1;
- The value of a heavy metal element concentration in any location is $\rho = z_1 - \Delta$.

With the help of concentration - Displacement offset model of Cd element, for example, we can plot the concentration of Cd in the geographical distribution as follows in Fig.2.

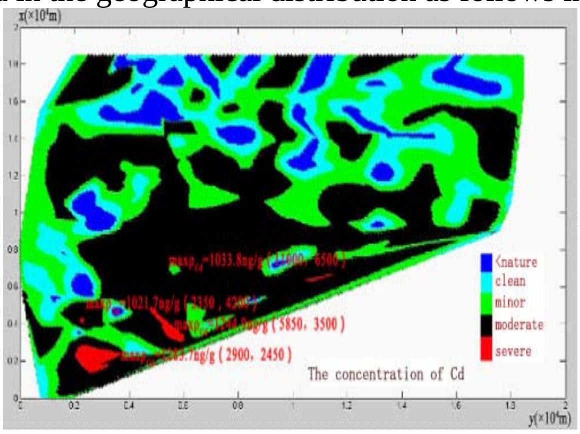

Fig.2 Concentration of Cd in the Geographical Distribution

As we can see from Fig.2 there are four pollution sources of Cd element, in which the concentration of pollution source a is 1021.7 ng/g and its position is (4200,2900,14.5513). The concentration of pollution source b is 1283.7 ng/g and its position is (2450,2900,14.5513). The concentration of pollution source c is 1244.9 ng/g and its position is (3550,5850,4.55270). The concentration of pollution source d is 1033.8 ng/g and its position is (6500,11000,2.4586). Severe contamination area represented in black in the figure occupies the largest size, followed by the moderate contamination area represented in green, furthermore blue and cyan composed of cleanness.

It can be concluded that pollution source a belongs to trunk road areas with a higher altitude; Areas near the pollution sources more than one kilometers radius are subjected to moderate contamination, the reason is that Cd element in the water flow leads to environmental damage[2]. We can conclude that the pollution sources may be the automotive recycling plant or gas station; Pollution source b with lower altitude also belongs to the trunk road areas and the periphery has a lot of green park space. This source 's influence on environment is relatively wide, which is mainly related to the physical characteristics of Cd element[5]. Under the influence of the precipitation, Cd element in the topsoil propagates horizontally with water, which results in over-concentration of Cd element in the green parks; The situation of pollution source c and pollution source b is similar, but size of contamination area caused by c is much larger than b, which is related to surroundings of c; Pollution source d belongs to the trunk road areas too, but because of in the hub of traffic arteries, the pollution caused by Cd element will be spread to a very broad range [6,7].

In summary, the propagation characteristics of Cd element are easy to migrate to a wider area with the diffusion of water and soil, and it's likely to be absorbed by plants.

Summary

In the analysis of the heavy metal pollution grade of different areas, we obtain the comprehensive pollution degree of each sampling point, namely cleanness, minor contamination, moderate contamination and severe contamination. Each sampling point has 8 indicators related with heavy metal. In order to get the weight of evaluation factor of each indicator, the fuzzy pattern recognition method is applied to evaluate the soil heavy metal pollution, avoiding subjective choosing the weight of each indicator and increasing the resolution of the evaluation results. Using the method to evaluate the urban soil, the results obtained reflect the actual situation of urban soils.

When describing the spatial distribution of important metals, if we only use the known sampling points, it merely can be able to draw some discrete data points, and it's unable to grasp the space height of the sampling points. With the help of the cubic interpolation method, we will not only depict the three-dimensional surface map, but also visually show the contamination degree for each sampling points with different colors.

In the model to find all pollution sources, the spatial distribution of the sampling points is discrete and uneven.By using "discrete matrix”,we transform it into a continuous problem, and at the same time, the concentration -displacement offset matrix is established.Therefore the relationship between concentration of each sampling point and its location will be established.Most important of all, the concise relationship is easy to calculate, and the error is smaller.

References

[1] Duan Xuemei, Cai Huanxing ,Chao Wenjun et al. Characteristics and sources of Nanjing surface soil heavy metal contamination [J],2010,10(5):31-34.

[2] Ren Yourong, *the complexity of the geological environment and disasters in city*[J], Geological Hazards and Environment,2003,14(3):1-4.

[3] Zheng Tong, Chen Chunyun et al. *Mathematical model in Environmental Systems* [M], Beijing: Chemical Industry Press,2003:110-136.

[4] Chen Yu, Zhou Lu, Qian et al. Numerical Methods [M], Beijing: Electronic Industry Press,2002,15(6):1-3.

[5] Fang Cunjin, *migration and transformation of heavy metal pollutants in soil* [J], contemporary chemical,2010,39(4):25-32.

[6] Information on http://baike.baidu.com/view/282970.htm.

[7] Zhang Zhihong, Yang Wenmin et al. *the analysis of chemical compositio Gasoline vehicles emit* ,2001,17(7):623-684.

Published by Atlantis Press, Paris, France.