第 12 卷 第 18 期 2012 年 6 月
1671—1815(2012)18-4461-07

科 学 技 术 与 工 程
Science Technology and Engineering

Vol. 12 No. 18 Jun. 2012
© 2012 Sci. Tech. Engrg.

环境科学

一种确定土壤重金属污染源位置的辨识方法

王明刚 许 华
(南京师范大学泰州学院,泰州 225300)

摘 要 确定土壤重金属污染源位置的问题是国内外研究的热门问题。考虑到当前城市土壤由于人类活动导致的环境剧烈变化,驱使自然生态系统下的自然成土过程向人工生态系统下的受人为因素干扰严重的特殊成土过程转变。提出利用因子分模型和灰色关联分析模型相结合的方法,确定污染源的位置。具体建模思路为:利用因子分析模型,筛选出受人为因素影响严重的重金属污染元素,然后利用灰色关联分析模型对筛选出的元素进行综合分析,进而确定污染源位置。

关键词 重金属污染 因子分析 灰色关联分析 变异系数

中图法分类号 X53; **文献标志码** A

城市土壤是城市生态环境的重要组成部分。对城市生态系统的稳定起着至关重要的作用。城市土壤由于人类活动导致的环境剧烈变化,驱使自然生态系统下的自然成土过程向人工生态系统下的受人为因素干扰严重的特殊成土过程转变。城市土壤重金属污染是城市土壤研究兴起的诱因,也是城市土壤研究的主要内容之一。城市土壤重金属分布在时间与空间上变异性明显。在较短距离内会出现显著变化,在水平方向上城市工业布局,产业结构的差异都会产生不同影响[1];在垂直方向上,不同功能区土壤剖面上层次排列混乱、分层明显,土壤重金属表层聚集和下沉现象兼而有之,外来物、侵入物、填埋物成份复杂更使差异增加。

在时间序列上由于城市化水平加速,产业结构调整,环境保护力度加大,新材料新工艺的使用也会使土壤重金属含量分布在较短时间内产生较大的变化。因为在城市不同的功能区人为活动强度各不相同,土壤重金属来源存在较大差异,污染程度和产生的危害也有所不同,因此,以上因素为研究城市地质环境的演变模式增加难度,但研究的重要性不容忽视。

确定土壤重金属污染源位置的问题是国内外研究的热门问题[2—6],但至今未有特别好的结果。考虑到当前城市土壤由于人类活动导致的环境剧烈变化驱使自然生态系统下的自然成土过程向人工生态系统下的受人为因素干扰严重的特殊成土过程转变,本文提出利用因子分析模型和灰色关联分析模型相结合的方法确定污染源的位置。具体建模思路为:利用因子分析模型筛选出受人为因素影响严重的重金属污染元素,然后利用灰色关联分析模型对筛选出的元素进行综合分析,进而确定污染源位置。

1 模型的建立

1.1 因子分析模型

因子分析是由英国心里学家 Spearman 在 1904 年提出来的,他成功解决了智力测验得分的统计分析[7],长期以来,教育心理学家不断丰富、发展了因子分析理论和方法,并应用这一方法在行为科学领域进行了广泛的研究。因子分析方法它根据多个实测变量之间的相互关系,运用数学变换,将多个变量转变为少数几个线性不相关的综合指标,从而

2012 年 3 月 16 日收到,3 月 23 日修改 江苏省高校大学生创新训练计划项目(2011126)、泰州市科技发展计划项目(2011045)、南京师范大学泰州学院院级项目资助
第一作者简介:王明刚(1982—),男,理学硕士,讲师,数学科学与应用学院院长助理,研究方向:数学建模,无穷维动力系统。E-mail: magic821204@ sina. com. cn。

简化数据处理,其目的在于对大量观测数据,用较少的有代表性的因子来说明众多变量所提取的主要信息,提示出多个变量间的因果关系。由于线性综合指标往往是不能直接观测到,但它更能反映事物的本质,因此因子分析在成因、来源问题上是一种非常有效的方法,可以很好地解决污染物来源的问题。

因子分析模型可以描述如下:设样本观察值构成的矩阵为

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \tag{1}$$

式(1)中,n 为样本观测数,p 为变量数(或称指标数),$X_j = (x_{1j}, x_{2j}, \cdots, x_{nj})^{\mathrm{T}}$,

$j = 1, 2, \cdots, p$。将原始数据按式(2)进行标准化处理。

$$x'_{ik} = (x_{ik} - \overline{x_k})/s_k; i = 1, 2, \cdots, n, k = 1, 2, \cdots, p \tag{2}$$

式(2)中

$$\overline{x_k} = \sum_{i=1}^{n} x_{ik}/n, s_k^2 = \sum_{i=1}^{n} (x_{ik} - \overline{x_k})^2/(n-1) \tag{3}$$

标准化处理后,变量的方差为 1,均值为 0。现将每个标准化处理后的变量用 $m(m < p)$ 个因子 $F_1, F_2, \cdots, F_m$ 的线性组合来表示,即有因子分析模型

$$\begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \varepsilon_1 \\ X_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + \varepsilon_2 \\ \vdots \\ X_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + \varepsilon_p \end{cases} \tag{4}$$

用矩阵表示为

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_P \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{pmatrix} \tag{5}$$

简记为

$$X = AF + \varepsilon \tag{6}$$

式(6)中,$X = (X_1, X_2, \cdots, X_p)^{\mathrm{T}}$ 为实测的 p 个指标所构成的 p 维向量;$A = (a_{ij})_{p\times m}$ 称为因子载荷矩阵,a_{ij} 为第 i 个变量在第 j 个因子上的载荷,$F = (F_1, F_2, \cdots, F_m)^{\mathrm{T}}$ 为不可观测向量,F 称为 X 的公共因子,$\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p)^{\mathrm{T}}$ 称为特殊因子。通常假定

(1) $m \leqslant p$;

(2) $\mathrm{cov}(F, \varepsilon) = 0$,即 F 与 ε 不相关;

(3) $\mathrm{var}(F) = I_m$, 即 $F_1, F_2, \cdots, F_m$ 不相关,且方差均为 1;

(4) $\mathrm{var}(\varepsilon) = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_p^2)$,

即 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p$ 不相关,且方差不同。

令 $h_i^2 = a_{i1}^2 + a_{i2}^2 + \cdots + a_{im}^2$,$h_i^2$ 称为共同度,它刻画了全部公共因子对变量 X_i 的总方差所做的贡献。如果对 X_i 已经进行了标准化,h_i^2 越接近于 1,说明该变量的全部原始信息被所选取的公共因子解释的就越多。

因子模型估计出来后,必须对所得的公共因子进行解释,若因子载荷矩阵中的系数居中,则将实施因子旋转,通常采用方差最大正交旋转,使得旋转后的因子载荷矩阵的每一列元素方差之和达到最大,且因子载荷矩阵中所有系数接近于 0 或 1,旋转后的因子载荷矩阵记为 A^*。利用旋转后的因子载荷矩阵,我们可以对重金属污染元素的来源进行区分,为建立灰色关联分析模型奠定基础。

1.2 灰色关联分析模型

灰色关联分析(Grey Relational Analysis)是灰色系统分析方法中的一种。基本上灰色关联分析是依据各因素数列曲线形状的接近程度做发展形势的分析。灰色系统理论提出了对各子系统进行灰色关联分析的概念,意图透过一定的方法,去寻求系统中各子系统(或因素)之间的数值关系。简言之,灰色关联分析的意义是指在系统发展过程中,如果两个因素变化的态势是一致的,即同步变化程度较高,则可以认为两者关联较大;反之,则两者关联较小。因此,灰色关联分析对于一个系统发展变化态势提供了量化的度量。

灰色关联动态分析的建模步骤如下[8]

(1)建立原始数列的因变量参考数列和自变量比较数列。

因变量参考数列又叫母序列记作

$X_0^{(k)}:X^{(k)} = [x_0^{(1)}, x_0^{(2)}, x_0^{(3)}, \cdots, x_0^{(k)}]$。

自变量比较数列又叫子序列记作

$X_i^{(k)}:X^{(k)} = [x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, \cdots, x_i^{(k)}]$,

$i = 1,2,\cdots,n$。

(2)将原始序列进行初值化法、均值化法的无量纲化处理,目的是消除数量级大小不同的影响,以便于进行计算和比较分析。

(3)计算每个时刻点上母序列与各子序列差的绝对值,并从中取得最大差和最小差。

差序列:$\Delta_i(k) = |x_0^{(k)} - x_i^{(k)}|, (i = 1,2,3,\cdots,n)$,

则差序列为

$\Delta_i = (\Delta_i(1), \Delta_i(2), \Delta_i(3), \cdots, \Delta_i(k))$,

$i = 1,2,3,\cdots,n$。

最大差:$\Delta_{\max} = \max_i \max_i |x_0^{(k)} - x_i^{(k)}|$;

最小差:$\Delta_{\min} = \min_i \min_i |x_0^{(k)} - x_i^{(k)}|$。

(4)计算灰色关联系数

$$L_{0i}^{(k)} = \frac{\Delta_{\min} + \lambda\Delta_{\max}}{\Delta + \lambda\Delta_{\max}}。$$

其中 $L_{0i}^{(k)}$ 是第 k 个点的子因素与母因素的相对差值。绝对差值 Δ 越大,$L_{0i}^{(k)}$ 越小,反之则 $L_{0i}^{(k)}$ 越大。因此 $L_{0i}^{(k)}$ 的大小描述了 X_i 对 X_0 的影响程度,称之为 X_i 与 X_0 在 k 处的点关联度。式中 λ 为分辨系数,一般在 0 与 1 之间取值,常取 $\lambda = 0.5$。

(5)计算灰色关联度。

为求总的关联度,需要考虑不同的观测点在总体观测中的重要性程度,则需要确定各点的权重,本文采用算术平均的方法计算灰色关联度 R_{0i}

$$R_{0i} = \frac{1}{n}\sum_{k=1}^{n} L_{0i}^{(k)}。$$

关联度越接近 1,说明关联程度越大。根据经验,当 $\lambda = 0.5$ 时,两因素的关联程度大于 0.6,便认为其关联性显著。因此,利用关联度的大小,可以区分受污染程度的高低。

2　实证分析

按照功能划分,城区一般可分为生活区、工业区、山区、主干道路区及公园绿地区等,分别记为 1 类区、2 类区、……、5 类区,不同的区域环境受人类活动影响的程度不同。

现对某城市城区土壤地质环境进行调查。为此,将所考察的城区划分为间距 1 km 左右的网格子区域,按照每平方公里 1 个采样点对表层土(0 ~ 10 cm 深度)进行取样、编号,并用 GPS 记录采样点的位置。应用专门仪器测试分析,获得了每个样本所含的多种化学元素的浓度数据。另一方面,按照 2 公里的间距在那些远离人群及工业活动的自然区取样,其作为该城区表层土壤中元素的背景值。利用 2011 年全国大学生数学建模竞赛 A 题所给数据[9],我们首先计算出八种重金属污染物的单因子指数和内梅罗综合指数(表 1)。

表 1　土壤重金属污染指数计算结果

功能区	单因子指数								内梅罗
	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn	综合指数
生活区	1.742	2.230	2.226	3.743	2.658	1.491	2.229	3.435	3.1707
工业区	2.014	3.024	1.723	9.662	18.353	1.611	3.001	4.028	13.5330
山区	1.122	1.172	1.257	1.312	1.170	1.256	1.179	1.062	1.2531
主干道路区	1.586	2.769	1.873	4.713	12.766	1.432	2.049	3.520	9.4262
公园绿地区	1.739	2.158	1.408	2.287	3.285	1.243	1.958	2.235	2.7340

为研究重金属污染物的相互关系,及寻找污染源,我们下面建立因子分析和灰色关联分析相结合的模型。

2.1　模型建立

2.1.1　对原始数据进行标准化处理

利用数据作为建模的原始数据,利用 Matlab 编程计算可得标准化指标向量为

$$X = \begin{pmatrix} 0.715\,4 & -0.660\,5 & -0.131\,4 & -0.211\,5 \\ 0.083\,8 & -0.694\,2 & -0.120\,8 & -0.199\,5 \\ -0.256\,8 & 0.608\,1 & -0.341\,9 & 0.058\,6 \\ 0.292\,1 & -0.348\,9 & -0.191\,8 & -0.183\,2 \\ 0.222\,7 & 0.990\,3 & 0.083\,4 & 0.383\,7 \\ 2.778\,7 & 3.513\,5 & 0.206\,4 & 1.556\,6 \\ 1.079\,1 & -0.144\,9 & 0.604\,6 & -0.062\,7 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$$\begin{pmatrix} -0.020\,7 & 0.094\,4 & -0.526\,6 & \cdots \\ -0.131\,1 & -0.0062 & -0.510\,6 & \cdots \\ -0.117\,0 & -0.670\,1 & 0.251\,3 & \cdots \\ 0.399\,1 & -0.187\,3 & -0.588\,5 & \cdots \\ 0.307\,0 & 0.295\,5 & 2.161\,9 & \cdots \\ 0.454\,3 & 1.100\,3 & 7.452\,6 & \cdots \\ -0.109\,7 & 0.054\,1 & 0.023\,4 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}。$$

2.1.2 计算相关系数矩阵 R

公式为：$R = (r_{ij})_{p\times p}$，$r_{ij} = \dfrac{\sum_{k=1}^{n} x_{ki}^{*} x_{kj}^{*}}{n-1}$；$(i,j = 1,2,\cdots,p)$，其中 $r_{ii} = 1$，$r_{ij} = r_{ji}$。

利用 Matlab 计算可得，该地区土壤重金属元素间的相关系数如表 2。

表 2 土壤重金属元素间的相关系数

元素	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn
As	1.000 0							
Cd	0.254 7	1.000 0						
Cr	0.189 0	0.352 4	1.000 0					
Cu	0.159 7	0.396 7	0.531 6	1.000 0				
Hg	0.064 4	0.264 7	0.103 2	0.416 7	1.000 0			
Ni	0.316 6	0.329 4	0.715 8	0.494 6	0.102 9	1.000 0		
Pb	0.289 9	0.660 3	0.382 8	0.520 0	0.298 1	0.306 8	1.000 0	
Zn	0.246 9	0.431 2	0.424 3	0.387 3	0.195 8	0.436 4	0.493 7	1.000 0

利用 Matlab 作聚类分析图如图 1。

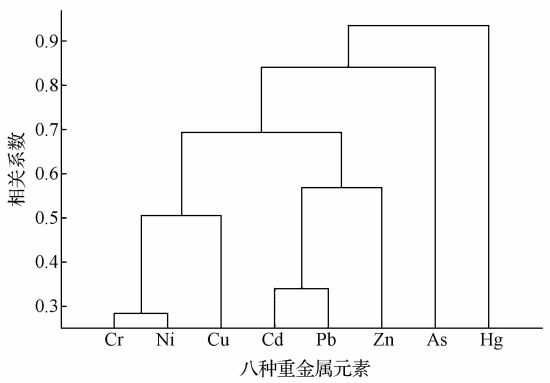

图 1 土壤重金属相关元素聚类分析图

可见 Cd 与 Pb，Cr、Cu 与 Ni，Cu 与 Pb 之间相关性显著，说明这些因素间可能具有同源关系或复合关系。

2.1.3 计算初等载荷矩阵

计算相关系数矩阵 R 的特征值 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p \geqslant 0$，及对应的特征向量 $u_1, u_2, \cdots, u_p$，其中 $u_j = (u_{1j}, u_{2j}, \cdots, u_{nj})^{\mathrm{T}}$，初等载荷矩阵 $A = [\sqrt{\lambda_1}u_1, \sqrt{\lambda_2}u_2, \cdots, \sqrt{\lambda_p}u_p]$。

带入数据计算得：

$$A = \begin{pmatrix} 0.425\,7 & -0.199\,6 & 0.681\,0 & 0.550\,8 & -0.026\,3 & -0.065\,0 & 0.007\,1 & 0.082\,3 \\ 0.710\,7 & 0.281\,4 & 0.282\,4 & -0.322\,1 & -0.254\,3 & 0.325\,0 & -0.220\,9 & 0.114\,3 \\ 0.735\,0 & -0.444\,0 & -0.303\,5 & -0.046\,2 & -0.109\,6 & 0.098\,2 & 0.285\,7 & 0.255\,2 \\ 0.756\,4 & 0.124\,7 & -0.365\,3 & 0.137\,5 & -0.154\,5 & -0.408\,1 & -0.250\,4 & 0.083\,0 \\ 0.408\,4 & 0.673\,4 & -0.297\,5 & 0.448\,7 & 0.154\,1 & 0.235\,6 & 0.1023 & -0.011\,7 \\ 0.722\,8 & -0.514\,6 & -0.189\,8 & 0.136\,7 & -0.013\,9 & 0.199\,7 & -0.119\,7 & -0.321\,9 \\ 0.764\,0 & 0.314\,3 & 0.237\,2 & -0.247\,8 & -0.157\,6 & -0.216\,7 & 0.283\,7 & -0.218\,0 \\ 0.698\,8 & -0.037\,4 & 0.123\,2 & -0.241\,0 & 0.653\,9 & -0.059\,8 & -0.055\,2 & 0.053\,5 \end{pmatrix}。$$

计算各个公共因子的贡献率如表 3，根据其累计贡献率的值，因为前 3 个主成份可以解释总变量的 70.9406%，因此对前 3 个主成份进行分析已经能够反映全部数据的大部分信息，我们选取前 3 个主因子。

表 3 旋转前因子贡献率

因子	贡献	贡献率	累计贡献率
1	3.56	44.499 9	44.499 9
2	1.150 2	14.377 4	58.877 3
3	0.965 1	12.063 3	70.940 6
4	0.767 7	9.596 1	80.536 7

2.1.4 进行因子旋转

对提取的因子载荷矩阵进行旋转，得到矩阵 $A^{*} = \tilde{A}T$，其中 $\tilde{A}$ 为 A 的前 3 列，T 为正交矩阵，利用 Matlab 程序计算得因子载荷矩阵

$A^{*} =$

$$\begin{pmatrix} 0.114\,7 & -0.202\,3 & 0.794\,2 & 0.550\,8 \\ 0.194\,0 & 0.462\,6 & 0.642\,2 & -0.322\,1 \\ 0.893\,8 & 0.101\,7 & 0.142\,1 & -0.046\,2 \\ 0.596\,9 & 0.590\,8 & 0.125\,1 & 0.137\,5 \\ 0.005\,3 & 0.841\,8 & 0.008\,2 & 0.448\,7 \\ 0.878\,6 & -0.001\,6 & 0.226\,7 & 0.136\,7 \\ 0.230\,1 & 0.530\,6 & 0.635\,8 & -0.247\,8 \\ 0.443\,9 & 0.258\,0 & 0.491\,2 & -0.241\,0 \\ -0.026\,3 & -0.065\,0 & 0.007\,1 & 0.082\,3 \\ -0.254\,3 & 0.325\,0 & -0.220\,9 & 0.114\,3 \\ -0.109\,6 & 0.098\,2 & 0.285\,7 & 0.255\,2 \\ -0.154\,5 & -0.408\,1 & -0.250\,4 & 0.083\,0 \\ 0.154\,1 & 0.235\,6 & 0.102\,3 & -0.011\,7 \\ -0.013\,9 & 0.199\,7 & -0.119\,7 & -0.321\,9 \\ -0.157\,6 & -0.216\,7 & 0.283\,7 & -0.218\,0 \\ 0.653\,9 & -0.059\,8 & -0.055\,2 & 0.053\,5 \end{pmatrix}。$$

旋转后的因子贡献及贡献率见表 4。

表 4　旋转后的因子贡献及贡献率

因子	贡献	贡献率	累计贡献率
1	2.227 9	27.85	27.85
2	1.671 2	20.89	48.74
3	1.776 1	22.20	70.94

表 5 给出因子旋转前后的载荷矩阵。

表 5　旋转前后的载荷矩阵

元素	旋转前			旋转后		
	因子 1	因子 2	因子 3	因子 1	因子 2	因子 3
As	0.425 7	-0.199 6	0.681 0	0.114 7	-0.202 3	0.794 2
Cd	0.710 7	0.281 4	0.282 4	0.194 0	0.462 6	0.642 2
Cr	0.735 0	-0.444	-0.303 5	0.893 8	0.101 7	0.142 1
Cu	0.756 4	0.124 7	-0.365 3	0.596 9	0.590 8	0.125 1
Hg	0.408 4	0.673 4	-0.297 5	0.005 3	0.841 8	0.008 2
Ni	0.722 8	-0.514 6	-0.189 8	0.878 6	-0.001 6	0.226 7
Pb	0.764 0	0.314 3	0.237 2	0.230 1	0.530 6	0.635 8
Zn	0.698 8	-0.037 4	0.123 2	0.443 9	0.258 0	0.491 2

利用 Matlab 编程绘出因子旋转前后因子得分散点图(图 2)。

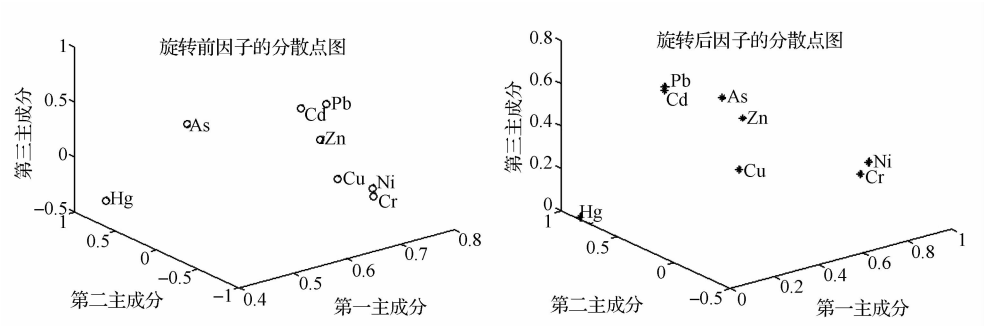

图 2　因子旋转前后因子的分散点图

2.2　结果解释

因子载荷矩阵旋转前每个因子在不同原始变量上的载荷没有明显差别,经旋转后载荷系数有了明显的两极分化。其中,第一主成分的贡献率为 27.85%,表现为因子变量在 Cr,Cu,Ni 上有较大的载荷,结合表 1 的这三种元素的污染指数,三种元素的平均污染指数分别为 1.697 4,4.343 4,1.406 6。而三种元素的变异系数相对较低,暗含了自然因子是控制它们的主要因素。该地区 Cu 污染较严重,这可能与该地区工业生产涉及 Cu、Cr 等金属配合物的催化剂的使用以及废水中富集的金属有关,同时 Cu 在第二主成分中也有较高载荷,暗示了 Cu 来源的多源性。Ni 主要受地球化学成因的影响,主要为地质来源,因此第一主成分基本反映了土壤母质及其风化产物而累计的重金属(自然来源)。

第二主成分的贡献率为 20.89%,在 Cu,Hg 和 Pb 上有较大的载荷,结合表 1 的这三种元素的污染指数,这三种金属元素的污染指数都较严重,污染指数平均值分别为 4.343 4,7.646 4,2.083 2,其中 Hg 污染最为严重,可能与工业三废”排放所致,Pb 的来源可能与道路交通有关,因此第二主成分基本反映了工业和交通造成的环境污染(人为来源)。

第三主成分的贡献率为 22.2%，在 As，Cd 和 Pb 上有较大的载荷，结合表 1 的这三种元素的污染指数，Cd 的平均污染指数 2.27 最高，这可能是由于该地区石油化工企业排放的氮氧化物和硫氧化物等有害气体，形成酸雨与地面重金属发生作用，使得周围的土壤容易富集高含量的有毒重金属。Cd 一般可作为施用农药和化肥等农业活动的标识元素，因此，第三主成分反映了农业生产等造成的环境污染。第二主成分与第三主成分合在一起反映了人类活动的影响，如施肥、工业活动产生的重金属颗粒物的干湿沉降、道路交通等对农田土壤重金属的贡献。

2.3 污染源位置的进一步确定

由以上的分析，我们仅需考虑受人为因素影响的重金属元素的污染情况，即考虑 As、Cd、Hg、Pb、Zn 的变异情况，由表 6 列出了 8 种重金属的变异系数。

表 6 8 种重金属的变异系数（单位：%）

重金属	生活区	工业区	山区	主干道路区	公园区
As	34.29	58.482 76	44.49	56.73	32.31
Cd	63.35	60.436 01	51.46	67.61	84.07
Cr	156.32	82.381 58	63.13	140.57	34.01
Cu	95.47	325.341 1	61.97	193.24	75.13
Hg	110.6	349.347 7	68.01	487.95	195.04
Ni	30.87	42.251 39	67.48	66.91	32.53
Pb	104.66	91.756 23	48.51	51.20	75.51
Zn	187.18	126.229 6	42.22	158.44	149.71

我们发现这 5 种重金属元素在工业区、主干道路区、生活区变异系数排在前三位，说明在这三个区人为因素影响显著，可以断定该区域的污染源主要来源于此三个区域。

下面以 As、Cd、Hg、Pb、Zn 五种重金属元素的浓度作为原始数据，以背景值作为参考数据列，建立灰色关联分析模型对样本点的关联度进行计算，部分结果见表 7。

结合题目所给出的采样点的位置坐标，我们可以在工业区、主干道路区和生活区这三个污染源的来源地区找到 8 个污染源，它们的位置坐标见表 8。

表 7 样本 5 种重金属污染的关联度（排序后）

编号	As/ (μg·g^{-1})	Cd/ (ng·g^{-1})	Hg/ (ng·g^{-1})	Pb/ (μg·g^{-1})	Zn/ (μg·g^{-1})	功能区	关联度
8	9.62	1 066.20	13 500.00	381.64	1 417.86	2	0.814 9
9	7.41	1 123.90	16 000.00	172.36	926.84	4	0.820 4
182	4.17	357.50	14 901.00	98.22	349.80	4	0.854 6
257	3.57	514.10	13 800.00	69.96	269.89	4	0.857 1
6	14.08	1 092.90	1 040.00	434.80	966.73	2	0.924 8
61	4.90	303.80	63.21	90.69	3 760.82	4	0.929 3
22	9.84	1 213.50	115.00	181.48	1 818.47	4	0.933 2
36	3.50	396.30	170.00	91.76	2 893.47	1	0.935 4
30	18.38	630.00	645.00	130.36	1 626.02	2	0.937 4
143	8.72	1 024.90	128.57	227.40	1 389.39	5	0.943 1
41	16.58	459.00	1 900.00	71.32	215.10	4	0.948
232	3.77	515.60	1 801.00	70.18	178.57	2	0.949 5
16	11.45	1 044.50	202.00	472.48	602.04	1	0.950 9
5	6.35	525.20	800.00	169.96	726.02	4	0.953 1
…	…	…	…	…	…	…	…

表 8 污染源的坐标（污染程度由大到小排序）

序号	x 坐标	y 坐标	z 坐标	来自区域
1	2 383	3 692	7	2
2	2 708	2 295	22	4
3	13 694	2 357	33	4
4	15 248	9 106	16	4
5	1 647	2 728	6	2
6	13 797	9 621	18	4
7	3 299	6 018	4	4
8	9 328	4 311	24	1

利用 Matlab 软件编程可以绘制出 8 个污染源的地理坐标（图 3）。

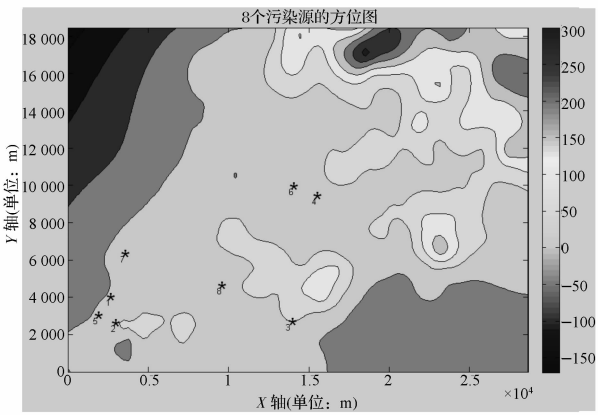

图 3 8 个污染源的方位图

3 总结

为了寻找最终的污染源，我们利用因子分析模

型与灰色关联分析模型相结合的方式达到了我们想要的目标。为了寻找污染源,我们利用因子分析模型和计算各重金属元素的变异系数首先判断污染源出自哪些人为因素,这是第一步缩小范围,然后再在小范围内进行寻找,相当于将整个空间进行分块。找到在哪一块后,我们再利用关联分析的方法进行定位,模型思路清晰,便于操作。

参 考 文 献

1 中华人民共和国国家环境保护局、国家技术监督局. GBl 5618—1995. 中华人民共和国国家标准. 北京:中国标准出版社,1996:03—01

2 Harrison R M, Laxen D P H, Wilson S J. Chemical association of lead, cadmium, copper and zinc in street dust and roads idesoil. Environ Sc i Techno l,1981;15:1378—1383

3 Thornton I. Metal contami nation of soils in urban areas. In: Bull ockP, Gregory P J. Soils in the Urban Environment. Blackwe l,1991:47— 75

4 段雪梅, 蔡焕兴, 巢文军. 南京市表层土壤重金属污染特征及污染来源. 环境科学与管理,2010;35(10):31—34

5 杨忠平,卢文喜,刘新荣,等. 长春市城区表层土壤重金属污染来源解析. 城市环境与城市生态,2009;22(5):29—33

6 王海东, 方凤满, 谢宏芳,等. 芜湖市区土壤重金属污染评价及来源分析. 城市环境与城市生态,2010;23(4):36—40

7 王吉权,邱立春,王富林,等. 引子分析法在高校教师教学质量评价中的应用. 数学的实践与认识,2011;41(15):28—34

8 刘思峰. 灰色系统理论及其应用(第三版). 北京:科学出版社,2004

9 教育部高等教育司. 2011 年高教社杯全国大学生数学建模竞赛赛题. http://www. mcm. edu. cn. 2011

Identification Method for Soil Heavy Metals Contaminants Source

WANG Ming-gang, XU Hua

(College of TaiZhou, Nanjing Normal University, Taizhou 225300, P. R. China)

[**Abstract**] Determine the location of soil heavy metal pollution has attracted wide attention. Taking into account the current urban soil environment due to human activities lead to dramatic changes in natural ecosystems driven into the soil under the natural process of the artificial ecosystem under serious interference by the special human factors into the process of changing the soil, the use of factor scores and gray correlation model analysis model to determine the method of combining the location of pollution sources are proposed. Specific modeling ideas are: the use of factor analysis model selected by the serious human factors elements of heavy metal pollution, and the use of gray correlation analysis model for screening out a comprehensive analysis of the elements, and to determine the location of pollution sources.

[**Key words**] heavy metal pollution　　factor analysis　　grey relational analysis　　variation coefficient