

2011 高教社杯全国大学生数学建模竞赛

承 诺 书

我们仔细阅读了中国大学生数学建模竞赛的竞赛规则.

我们完全明白, 在竞赛开始后参赛队员不能以任何方式(包括电话、电子邮件、网上咨询等)与队外的任何人(包括指导教师)研究、讨论与赛题有关的问题。

我们知道, 抄袭别人的成果是违反竞赛规则的, 如果引用别人的成果或其他公开的资料(包括网上查到的资料), 必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺, 严格遵守竞赛规则, 以保证竞赛的公正、公平性。如有违反竞赛规则的行为, 我们将受到严肃处理。

我们参赛选择的题号是(从 A/B/C/D 中选择一项填写): A

我们的参赛报名号为(如果赛区设置报名号的话): 13212002

所属学校(请填写完整的全名): 湖北师范学院

参赛队员(打印并签名): 1. 易汀涓

2. 刘慧娜

3. 李端君

指导教师或指导教师组负责人(打印并签名): 陈琴

日期: 2011 年 9 月 12 日

赛区评阅编号(由赛区组委会评阅前进行编号):

2011 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅编号（由赛区组委会评阅前进行编号）：

赛区评阅记录（可供赛区评阅时使用）：

评 阅 人										
评 分										
备 注										

全国统一编号（由赛区组委会送交全国前编号）：

全国评阅编号（由全国组委会评阅前进行编号）：

城市表层土壤重金属污染分析

摘要

本文就重金属在某城市五类区域的浓度分布,利用 Surfer 软件绘出了八种重金属在该城区的分布图,我们保持单一变量法,对每种元素在整个区域的分布情况,利用等高线思想汇出图 1 至图 8。

问题一根据单因子指数 $P_i = \frac{1}{m} \sum_{i=1}^m C_i / S_i$ 和内梅罗指数 $P_{\text{综}} = \sqrt{\frac{\left(\frac{1}{n} \sum_{i=1}^n P_i\right)^2 + (\max P_i)^2}{2}}$

(C_i 和 S_i 分别表示污染物的测量值和背景值) 确定五个城区的金属污染程度。并利用

$\omega_i = \frac{\gamma_i}{\sum_{i=1}^5 \gamma_i}$ ($\gamma_i = \frac{P_{\text{max}}}{P_i}$ 表示每种污染物的相对重要性比值, ω_i 为各评价指标权重) 对内

梅罗指数进行了加权平均修正, 得出:

区域名称	修正综合污染指数	污染等级
生活区	2.569	中度污染
工业区	8.374	重度污染
山区	1.214	轻度污染
交通区	5.784	重度污染
公园绿地区	2.227	中度污染

在问题二中, 本文先进行横纵比较, 初步得工业区 Hg 的污染程度最严重, 然后建立主成分分析模型, 利用模型一中的单因子指数, 建立矩阵, 利用 matlab 进行处理后, 得到相关系数矩阵, 计算的八个特征值分别为: 6.3757, 0.9437, 0.4392, 0.2415, 0, 0, 0, 0。据累计贡献率得出三个主成分, 由此确定工业区 Hg、Cu, 交通区 As、Hg, 生活区 Cu 为重金属污染的主要原因。

问题三从三个侧面进行数据分析, 判断重金属污染物的三个传播特征: 以污染源为圆心逐步向外扩散; 以工业区为主要传播重心, 随以重金属富集明显的路径进行传播; 某些重金属不会扩散, 其在土壤的稳定性很强, 在一定时期内不会传播。参考地震源的位置确定采用 Geiger 定位方法, 考虑单一污染物的浓度分布情况, 利用 Inglada 方法进行

初值的计算得 $\frac{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2}{(C_i - C_0)} = \frac{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2}{(C_j - C_0)}$

((x_0, y_0, z_0) 和 (x_i, y_i, z_i) 分别为污染源、观测点坐标, C_i, C_0 分别为观测点、污染源浓度)。

利用 Geiger 方法迭代校正, 确定 Cu 污染源位置在 $0 \leq x_0 \leq 2454, 987 \leq y_0 \leq 1570$ 的范围。

在问题四中, 我们充分考虑了模型三的静态特征, 还应该收集土壤的固相、液相、气相和 PH 值以及重金属元素的毒理性的相关信息。此外, 我们还化静态为动态加入时间参变量, 提高了模型三的准确度。整合以上信息, 可以利用差分方程建立模型求解。

关键词: 单因子指数 内梅罗指数 主成分分析法 Geiger 定位法 Inglada 方法

一、问题重述

随着城市经济的快速发展和城市人口的不断增加，人类活动对城市环境质量的影响日显突出。对城市土壤地质环境异常的查证，以及如何应用查证获得的海量数据资料开展城市环境质量评价，研究人类活动影响下城市地质环境的演变模式，日益成为人们关注的焦点。

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 公里左右的网格子区域，按照每平方公里 1 个采样点对表层土（0~10 厘米深度）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，将其作为该城区表层土壤中元素的背景值。

附件 1 列出了采样点的位置、海拔高度及其所属功能区等信息，附件 2 列出了 8 种主要重金属元素在采样点处的浓度，附件 3 列出了 8 种主要重金属元素的背景值。

现要求你们通过数学建模来完成以下任务：

- (1) 给出 8 种主要重金属元素在该城区的空间分布，并分析该城区内不同区域重金属的污染程度。
- (2) 通过数据分析，说明重金属污染的主要原因。
- (3) 分析重金属污染物的传播特征，由此建立模型，确定污染源的位置。
- (4) 分析所建立模型的优缺点，为更好地研究城市地质环境的演变模式，还应收集什么信息？有了这些信息，如何建立模型解决问题？

二、问题分析

本问题是关于环境生物和地址学相结合的模型问题，附件给出了区域、采样点的坐标位置、海拔以及污染物浓度一系列的数据。四个问题从不同的侧面要求我们讨论污染程度、主要原因、传播特征以及污染源的位置确定，最终是希望我们客观全面的研究地质环境的演变模式，从而有效地解决环境污染问题。

问题一：基于附件 1 中采样点的位置和附件 2 中 8 种重金属元素在各个区域内的分布情况，运用 `surfer` 软件绘出下列 8 个等浓度图。考虑重金属在该城区的空间分布情况时，我们针对每种元素对不同的 5 个区域进行了绘制，在分析该城区重金属的污染程度时，我们针对每类区域的 8 种重金属的浓度，采用单因子指数法和内梅罗综合污染指数法进行污染程度的分析。并且在此基础上，我们对内梅罗指数进行了修正，得出比较客观的修正值，让结果变得更加准确。

问题二：首先进行横向比较和纵向比较，初步进行直观判断，然后进行主成分分析，得出三个主成分，和初步分析的结果基本吻合。

问题三：分别从不同侧面进行分析，即：纵横坐标、海拔、区域以及重金属污染物固有的一些特性，然后跟确定地震源的确定方法进行比较和借鉴，我们采用了经典的 Geiger 定位法，并且还运用 Inglada 方法进行初值的计算。然后再用 Geiger 进行迭代校

正。这样是结果更加精准。

问题四：这一问属于开放性问题，所以在模型三的基础上，我们只要加上相关的参数，并且将模型三有静态改为动态模型，这样我们加上时间参变量，对其进行模型改造。利用差分方程进行有效的解决。

三、模型假设

- 1、假设每类区表层土壤中每种重金属含量连续变化；
- 2、在对采样点分析时，忽略土壤 PH 值、温度、气候等因素对测量带来的影响；
- 3、在对每类区域的污染程度进行分析的时候，忽略考察的 8 种元素以外的重金属污染；
- 4、忽略重金属由于土壤特性和氧化还原等一系列的物理和化学作用，给污染程度评价带来的误差。
- 5、由于不好判断五类区的地理分布情况，我们就将问题理想化，考虑每类区的 8 种重金属污染程度；
- 6、假设附件给出的数据对于建立模型已经足够充分；
- 7、假设每个采样点的采集时间相同，重金属的含量不会因为采集时间的差异出现变化；
- 8、忽略污染化合物之间的相关影响；
- 9、忽略重金属污染物之间的交叉污染；
- 10、在第三问中，我们将土壤理想化成成分均匀的介质。

四、符号说明

P_i : 土壤重金属污染物 i 的环境质量指数

m : 土壤重金属污染物 i 采样点的个数

C_i : 土壤重金属污染物 i 的实际测量值

S_i : 土壤重金属污染物 i 的背景值

n : 本文重金属污染物种类总数

$P_{\text{综}}$: 综合环境质量指数

γ_i : 相对重要性比值

ω_i : 各评价指标的权重

$P_{\text{修}}$: 综合环境质量指数修正值

(x_0, y_0, z_0) : 污染源坐标

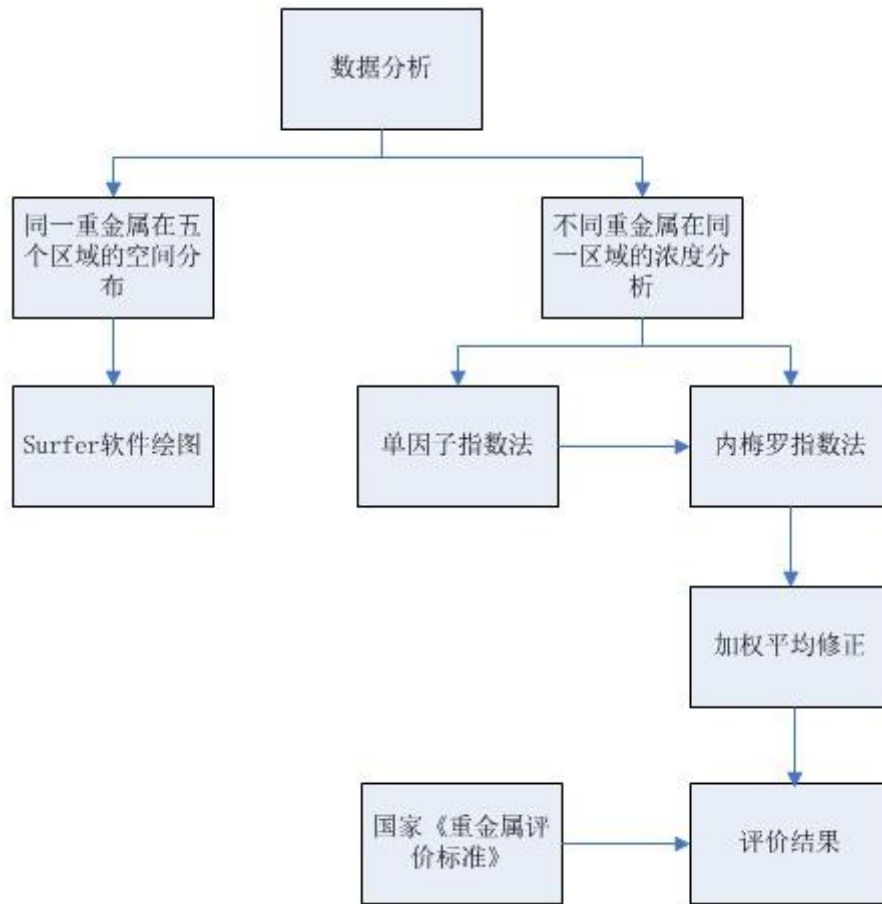
$\Delta x, \Delta y, \Delta z$: 污染源坐标校正量

ΔC : 浓度校正量

r_k : 测定浓度的偏差

五、模型的建立与求解

5.1 问题一



5.1.1 八种重金属元素的空间分布

在考虑 8 种重金属在该城区的空间分布时，本文抓住问题的核心：单一元素在五类区域的空间分布。这样我们利用附件一和附件二使用 matlab 软件绘出 8 种重金属在整个区域的空间分布等高线（见图 1-8）。

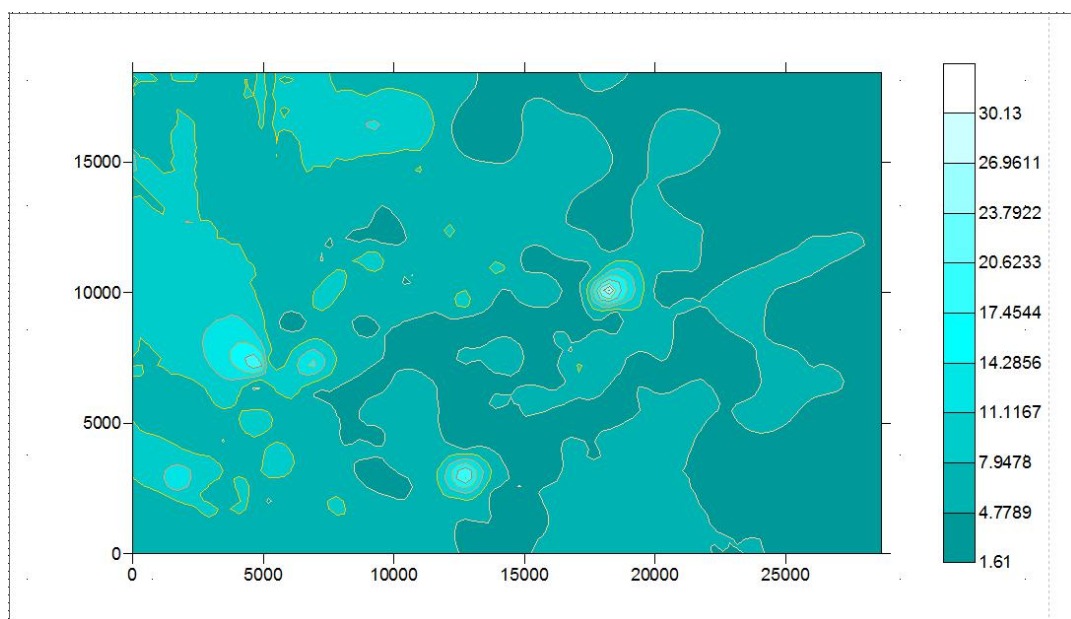


图 1 重金属元素 As 的空间分布图

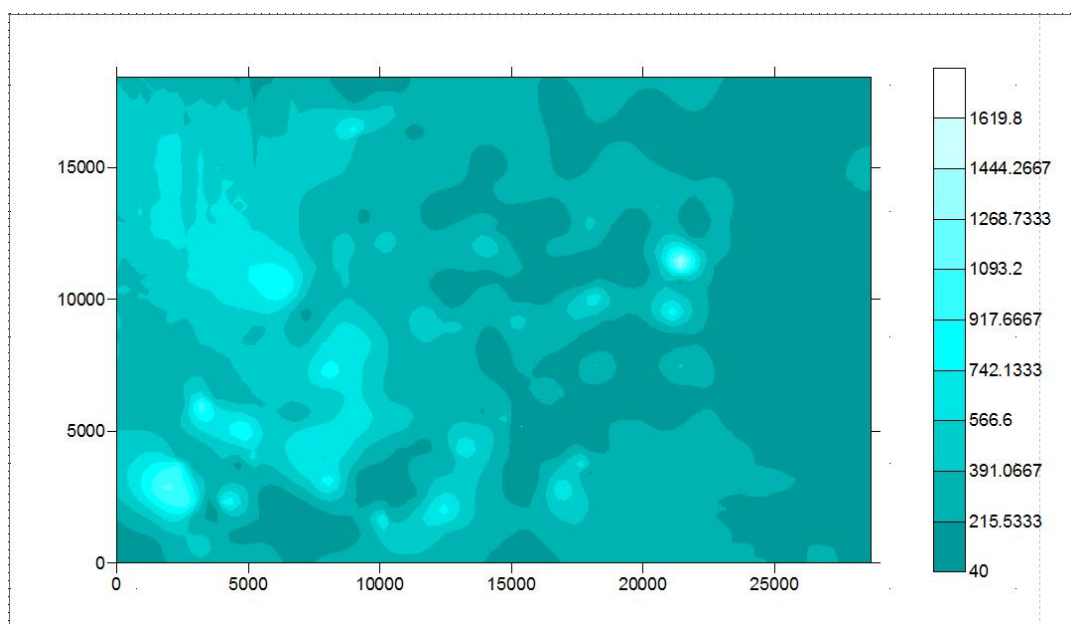


图 2 重金属元素 Cd 的空间分布图

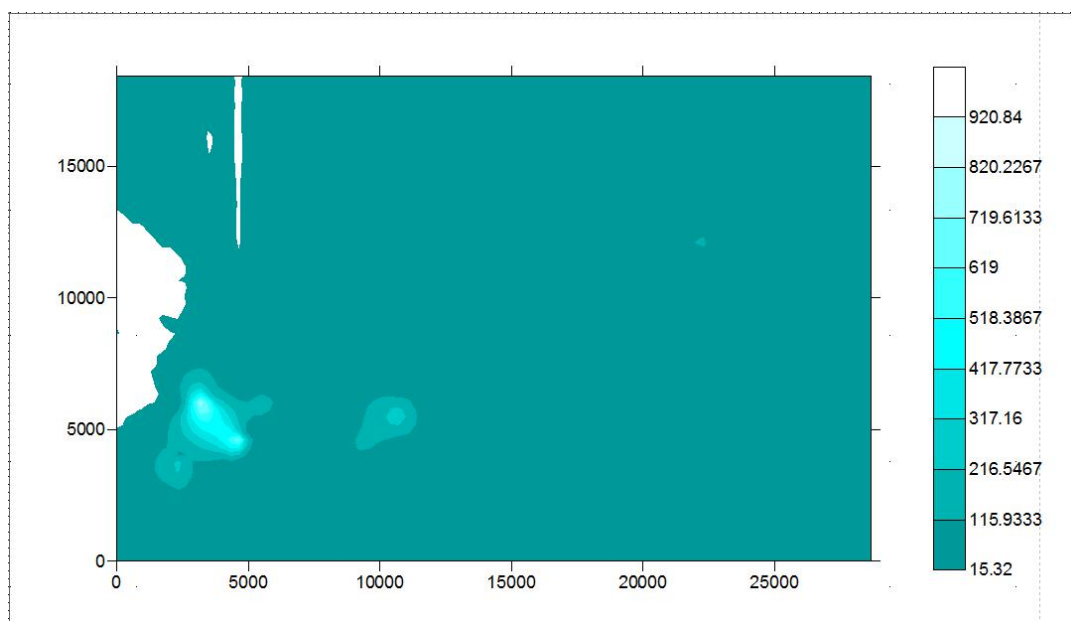


图 3 重金属元素 Cr 的空间分布图

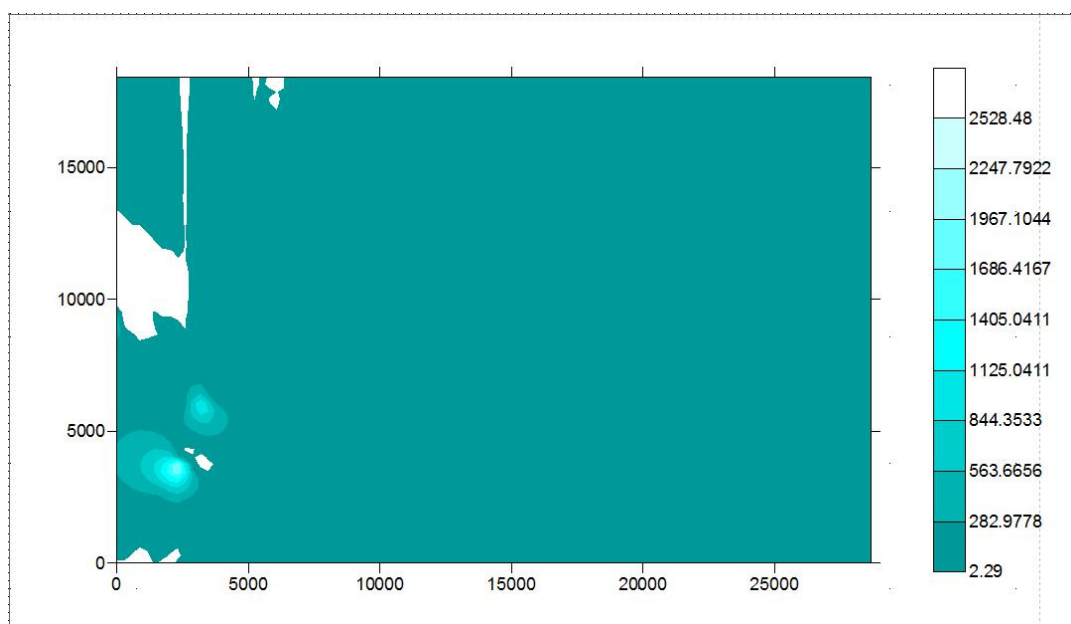


图 4 重金属元素 Cu 的空间分布图

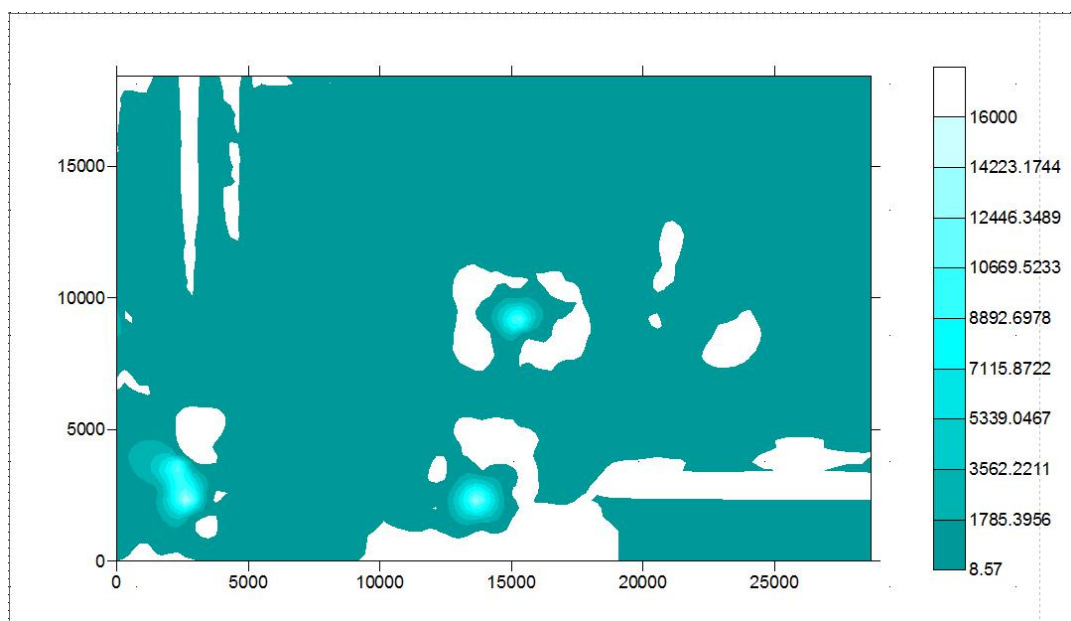


图 5 重金属元素 Hg 的空间分布图

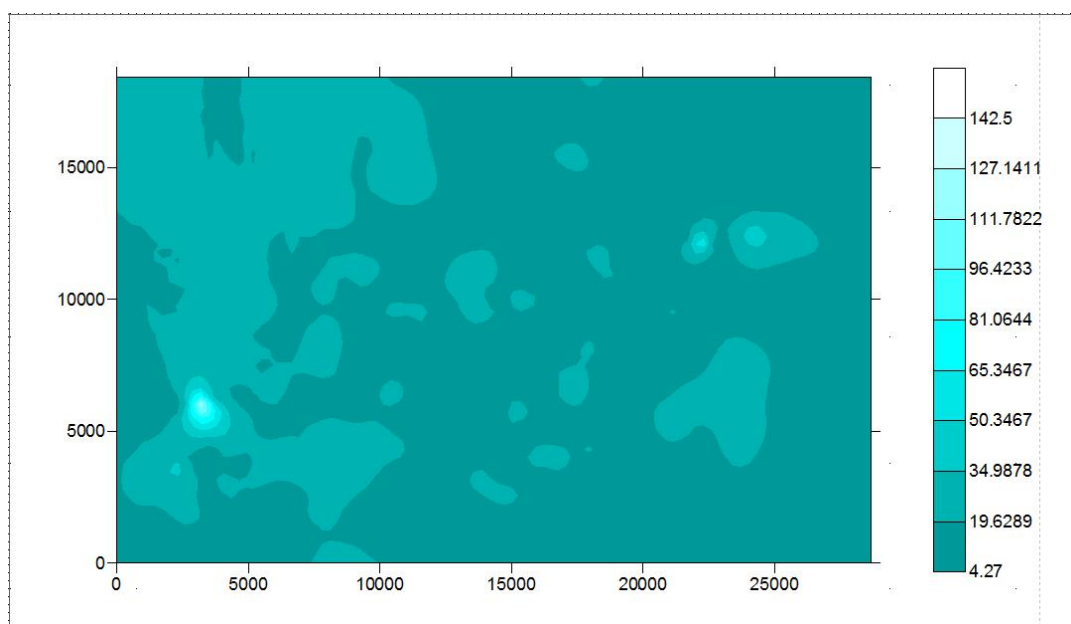


图 6 重金属元素 Ni 的空间分布图

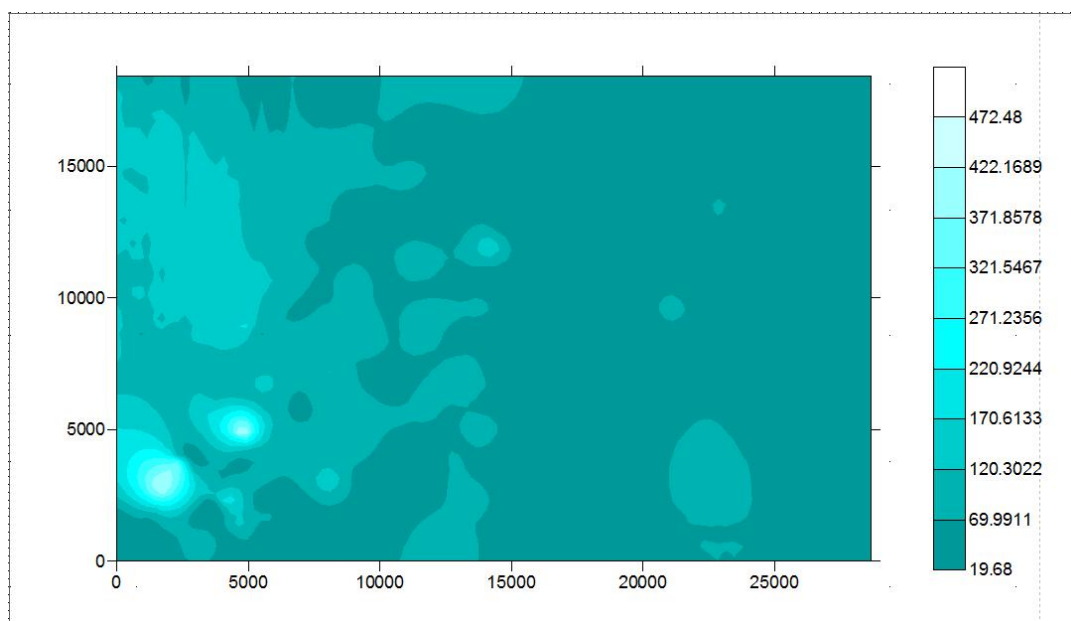


图 7 重金属元素 Pb 的空间分布图

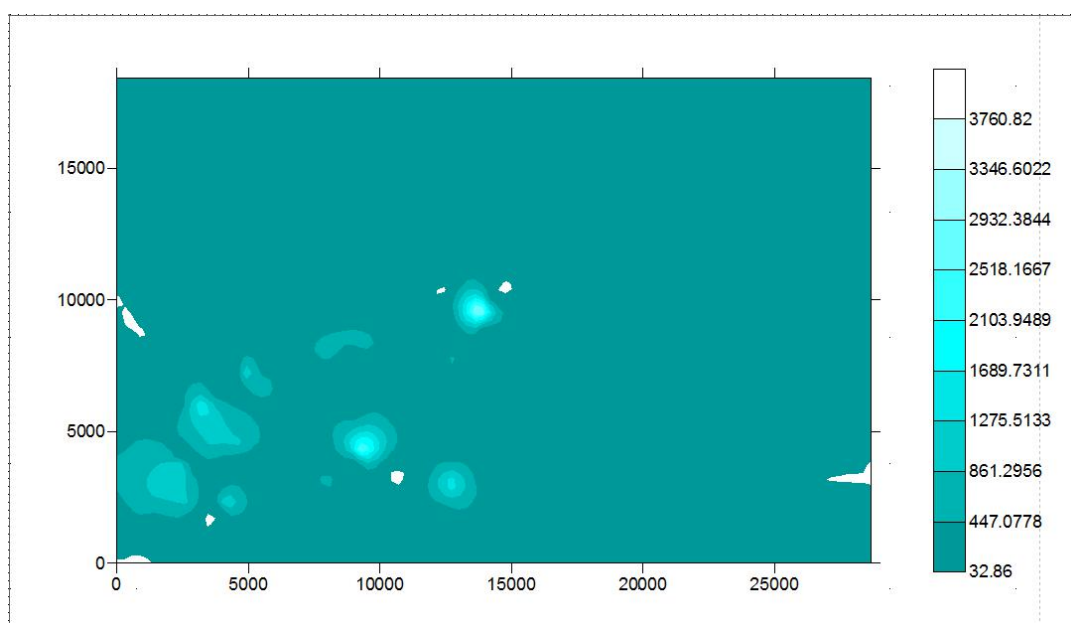


图 8 重金属元素 Zn 的空间分布图

5.1.2 不同地区重金属的污染程度

在分析该城区内不同区域重金属的污染程度时，本文使用单因子指数法和综合因子指数法。

1、评价方法

首先采用单因子污染指数法对土壤重金属污染进行评价，计算公式为：

$$P_i = \frac{1}{m} \sum_{i=1}^m C_i / S_i$$

然后采用综合因子指数法对土壤重金属污染进行评价，计算公式为：

$$P_{\text{综}} = \sqrt{\frac{\left(\frac{1}{n} \sum_{i=1}^n P_i\right)^2 + (\max P_i)^2}{2}}$$

式中， P_i 为土壤污染物 i 的环境质量指数； C_i 为土壤污染物 i 的实际测量值； S_i 为土壤污染物 i 的背景值； m 是土壤污染物 i 采样点的个数； n 为重金属种类总数。

2、评价标准

采用国家《土壤重金属评价标准》(GB15618-1995)^[1]进行评价，评价标准见表 1、2：

表 1 土壤各元素污染程度分级标准

等级划分	1	2	3	4
单因子污染指数	$P_i \leq 1$	$1 < P_i \leq 2$	$2 < P_i \leq 3$	$P_i \geq 3$
污染程度	非污染	轻度污染	中度污染	重度污染

表 2 土壤综合污染程度分级标准

等级划分	1	2	3	4	5
综合污染指数	$P_{\text{综}} \leq 0.7$	$0.7 < P_{\text{综}} \leq 1$	$1 < P_{\text{综}} \leq 2$	$2 < P_{\text{综}} \leq 3$	$P_{\text{综}} \geq 3$
污染程度	安全	警戒线	轻度污染	中度污染	重度污染

3、评价结果

使用 EXCEL 进行数据分析与处理，得到如下评价结果：

表 3. 不同区域重金属的污染指数

区域	单因子污染指数 P_i								综合污染指数 $P_{\text{综}}$
	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn	
1	1.74	2.23	2.23	3.74	2.66	1.49	2.23	3.43	3.17
2	2.01	3.02	1.72	9.66	18.35	1.61	3.00	4.03	13.53
3	1.12	1.17	1.26	1.31	1.17	1.26	1.18	1.06	1.25
4	1.59	2.77	1.87	4.71	12.77	1.43	2.05	3.52	9.43
5	1.74	2.16	1.41	2.29	3.29	1.24	1.96	2.24	2.73

单因子污染指数评价结果：生活区中，Cu 和 Zn 属于重度污染；工业区中，Hg 和 Cu 的污染程度及其严重；山区各重金属污染程度均较低；交通区中也是 Hg 的污染相当严重；公园绿地区各重金属污染程度相对还好。

5.1.3 修正内梅罗指数法

显然，用此方法的计算结果会突出最大污染物产生的影响，但却没有考虑污染因子的危害性差异^[2]。会引入很大的误差，因此，将此式修改如下：

将污染指数平均值改为加权平均，根据评价指标对环境及人体的危害性来确定各评价指标的权重，具体做法如下：

首先将综合评价指标按由小到大的顺序排列为 $P_{\text{综}1}, P_{\text{综}2} \cdots P_{\text{综}n}$ 将最大的一个记为

P_{\max} ，令 $\gamma_i = \frac{P_{\max}}{P_i}$ 表示第 i 种评价指标的相对重要性比值，则 $\omega_i = \frac{\gamma_i}{\sum_{i=1}^5 \gamma_i}$ 为各评价指标的

权重。下面运用 Excel 进行简单计算，结果如下（表 4）：

表 4 修正后的综合污染指数

区域名	综合指数	相对重要性比值 γ_i	各评价指标 的权重 ω_i	修正值 $P_{\text{修}}$
生活区	3.17	0.221045	0.099919	2.569
工业区	13.53	1	0.45203	8.374
山区	1.25	0.092596	0.041856	1.214
交通区	9.43	0.696549	0.314861	5.784
公园绿地区	2.73	0.202054	0.091335	2.227

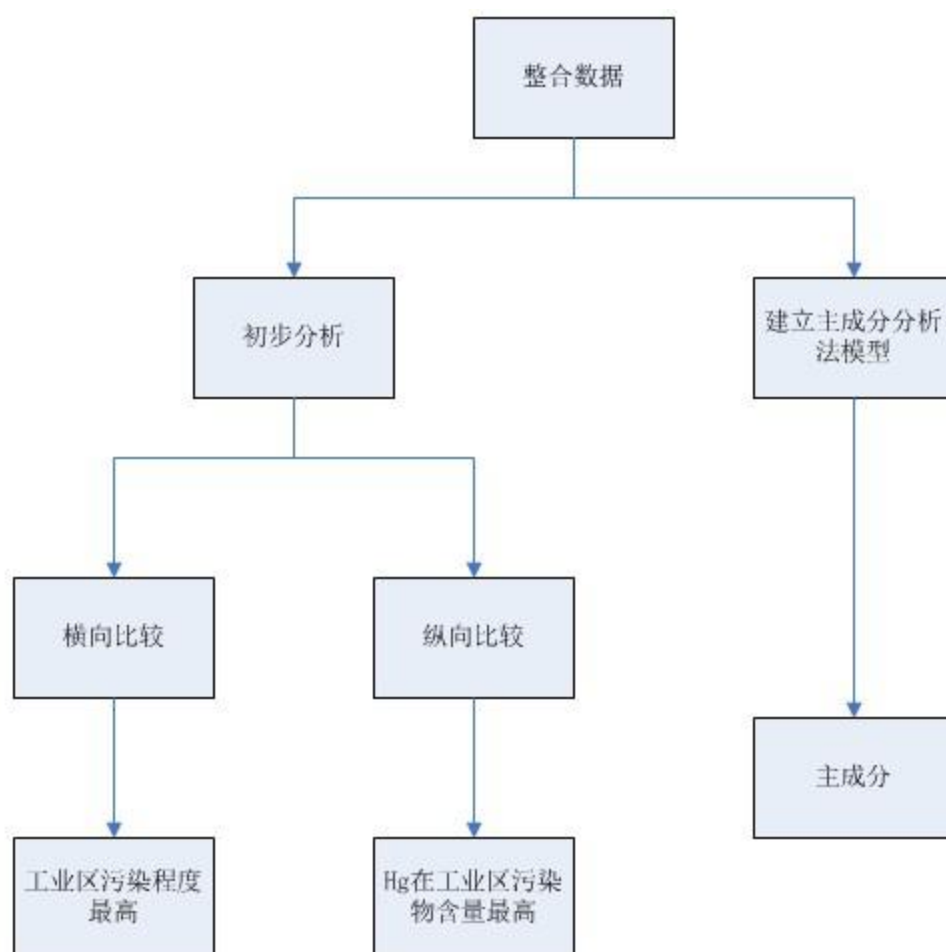
修正后得到不同城区的综合污染程度（表 5）。

表 5 不同城区综合污染程度

区域名称	综合污染指数	污染等级
生活区	2.569	中度污染
工业区	8.374	重度污染
山区	1.214	轻度污染
交通区	5.784	重度污染
公园绿地区	2.227	中度污染

综合污染指数评价结果：工业区污染指数为 8.374，属于重度污染，其次是交通区，污染指数为 5.784，生活区和公园绿地区属于中度污染，污染指数分别为 2.569 和 2.227，山区属于轻度污染，污染指数为 1.214。

5.2 问题二



5.2.1 初步分析

通过问题一中对于重金属污染的数据表 3 和表 5,我们分别从横向和纵向进行分析,找出重金属污染的主要原因。

I. 横向比较

横向比较五类综合污染指数,可以清晰地看出,工业区所占比例明显高出其他四类区域,交通区次之,生活区和公园绿化区持平居后,山区的污染程度最轻。这说明工业污染导致重金属污染的成分最重,交通区主要以排放的污染气体为主,被污染的大气长期滞留,这样就会污染到土壤表层。对于生活区和公园绿化区,分析得知,他们的日常生活基本一致,然而公园绿化区的污染程度要比生活区稍轻一些,这跟实际情况是相吻合的。下图给出了五类区的污染指数所占比例:

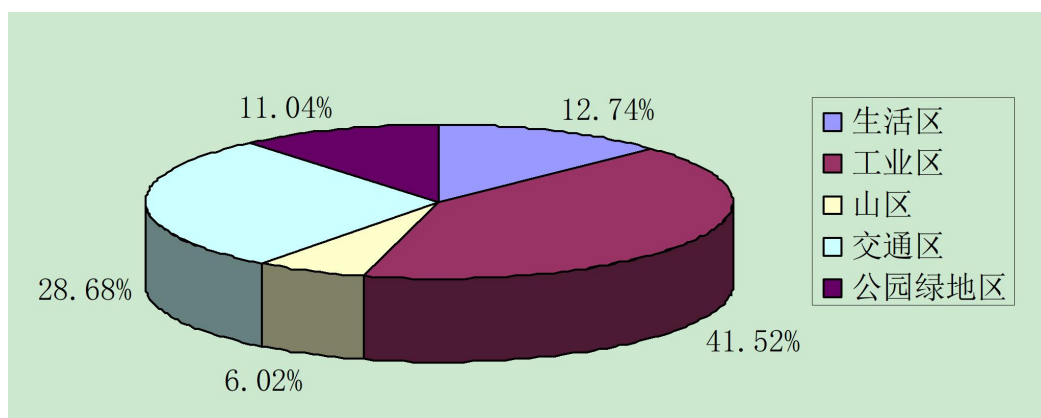


图9 五类区域综合污染指数所占比例

横向分析可以看出，工业区的综合指数最高，这与实际规律相符，说明工业区周围的治理污染工作还不到位，横向主要原因为工业污染。

II. 纵向比较

纵向比较污染程度最严重的区域（工业区），分析8种重金属的单因子指数，这样可以明显找出导致污染的主要因素。

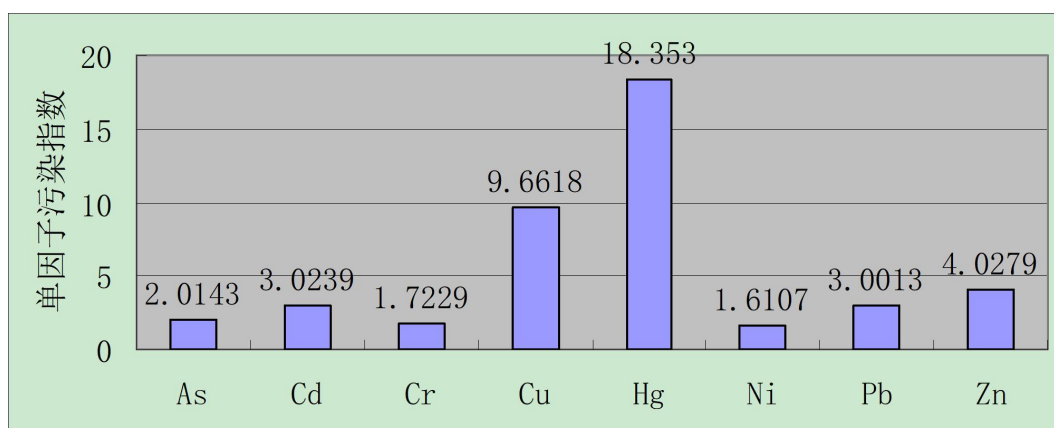


图10 工业区单因子污染指数

由图10可以看出Hg的污染指数最高，工业区土壤中Hg明显富集，这与工业区的Hg元素及含Hg的化合物渗透土壤的污染物比较多，下面我们采用主成分分析法进行污染物的主要原因的确定。

5.2.2 主成分分析

先利用问题一处理的结果，建立五个区域和八种元素的五行八列的单因子矩阵 X

$$X = \begin{pmatrix} 1.74 & 2.23 & 2.23 & 3.74 & 2.66 & 1.49 & 2.23 & 3.43 \\ 2.01 & 3.02 & 1.72 & 9.66 & 18.35 & 1.61 & 3.00 & 4.03 \\ 1.12 & 1.17 & 1.26 & 1.31 & 1.17 & 1.26 & 1.18 & 1.06 \\ 1.59 & 2.77 & 1.87 & 4.71 & 12.77 & 1.43 & 2.05 & 3.52 \\ 1.74 & 2.16 & 1.41 & 2.29 & 3.29 & 1.24 & 1.96 & 2.24 \end{pmatrix}$$

然后利用 matlab 计算 X 的相关系数方阵 R

$$R = \begin{pmatrix} 1.0000 & 0.8653 & 0.5007 & 0.7657 & 0.6219 & 0.6915 & 0.9526 & 0.8463 \\ 0.8653 & 1.0000 & 0.5436 & 0.8370 & 0.8562 & 0.7618 & 0.9003 & 0.9445 \\ 0.5007 & 0.5436 & 1.0000 & 0.3572 & 0.2214 & 0.6873 & 0.5247 & 0.7664 \\ 0.7657 & 0.8370 & 0.3572 & 1.0000 & 0.9245 & 0.8978 & 0.9187 & 0.8249 \\ 0.6219 & 0.8562 & 0.2214 & 0.9245 & 1.0000 & 0.7542 & 0.7807 & 0.7594 \\ 0.6915 & 0.7618 & 0.6873 & 0.8978 & 0.7542 & 1.0000 & 0.8606 & 0.8892 \\ 0.9526 & 0.9003 & 0.5247 & 0.9187 & 0.7807 & 0.8606 & 1.0000 & 0.9085 \\ 0.8463 & 0.9445 & 0.7664 & 0.8249 & 0.7594 & 0.8892 & 0.9085 & 1.0000 \end{pmatrix}$$

其中 r_{ij} 为 R 中 i 行 j 列的元素，其计算公式为

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$$

由此公式可以知道 $r_{ij} = r_{ji}$ ，然后计算上述方阵的特征值 λ ，令 $|\lambda I - R| = 0$ ，由 matlab 计算结果为：

然后进行排序 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ，这样得到特征向量 e_i ， $\sum_{j=1}^p e_{ij}^2 = 1$ 得到贡献率：

$$\frac{\lambda_i}{\sum_{k=1}^p \lambda_k}, \text{ 累计贡献率: } \lambda'_i = \frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k}, \text{ 利用 matlab 计算结果见下表: (表 6)}$$

表 6 特征值、贡献率、累计贡献率

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
特征值	6.3757	0.9437	0.4392	0.2415	0.0000	0.0000	0.0000	0.0000
贡献率	0.7970	0.1180	0.0549	0.0302	0.0000	0.0000	0.0000	0.0000
累计贡献率	0.7970	0.9149	0.9698	1.0000	1.0000	1.0000	1.0000	1.0000

表 7 特征向量

0.1519	0.3559	0.4057	-0.2072	-0.2221	-0.6861	0.0121	0.3500
0.4885	-0.1233	-0.1868	0.4426	0.5757	-0.1959	-0.0681	0.3755
-0.2844	0.1362	0.2751	0.2991	0.1234	0.1804	0.7914	0.2476
-0.0309	-0.5368	0.5147	0.2521	-0.3149	0.2269	-0.3107	0.3678
-0.3643	0.5061	0.1305	-0.1097	0.3837	0.3307	-0.4619	0.3349
0.5209	0.2952	-0.2610	-0.1244	-0.4283	0.4779	0.1182	0.3635
-0.4955	-0.0429	-0.5994	0.2553	-0.3274	-0.2628	-0.0680	0.3839
-0.0829	-0.4541	-0.1248	-0.7180	0.2589	0.0053	0.2011	0.3851

取累计贡献率达 85%—95%的特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ 所对应的第一、第二…第 $m(m \leq p)$ 的主成分，那么就得到主成分有三个，分别为第一主成分 F_1 第二主成分 F_2 第三主成分 F_3 ：

$$F_1 = 0.3851x_1 + 0.2011x_2 + 0.00538x_3 + 0.25894x_4 - 0.7180x_5 - 0.1248x_6 - 0.4541x_7 - 0.0829x_8$$

$$F_2 = 0.3839x_1 - 0.0680x_2 - 0.2628x_3 - 0.3274x_4 - 0.2553x_5 - 0.5994x_6 - 0.0429x_7 - 0.4955x_8$$

$$F_3 = 0.3635x_1 + 0.1182x_2 + 0.4779x_3 - 0.4283x_4 - 0.1244x_5 - 0.2610x_6 + 0.2952x_7 + 0.5209x_8$$

由此确定污染主要原因为：工业区 Hg、Cu，交通区 As、Hg，生活区 Cu，分别为重金属污染的主要原因。

5.3 问题三

5.3.1 重金属污染物的传播特征

对于重金属污染物的空间分布，结合数据的变化情况，我们首先从几个侧面分析污染物的传播特征，然后根据最速下降法进行方向追踪污染源的位置。

我们整合了题目给的附表一和附表二的数据，建立了表一（见附录），有数据直观分析出，从平面角度：以污染物浓度较高的地理位置一圆心扩散，从区域划分角度：工业区——>交通区——>生活区——>公园绿化区——>山区。然而，值得注意的是，从元素的种类上分析时，我们发现，对于某些重金属不会扩散，也就是说，其在土壤的稳定性很强，在一定时期内不会传播，那么我们在确定传染源的时候，就进行模糊聚类，将一些传播明显的元素进行梯度分析，进而得到最速下降最快的方向，也就得到了污染源的位置。

5.3.2 土壤重金属污染来源介绍

土壤中重金属元素主要有自然来源和人为干扰输入两种途径。^[3]在自然因素中，成土母质和成土过程对土壤重金属含量的影响很大。在各种人为因素中，则主要包括工业和交通等来源引起的土壤重金属污染。以下主要就受人为作用影响的土壤重金属污染来源进行介绍。

1、不同工矿企业对重金属积累的影响

工业过程中广泛使用重金属元素，工矿企业将未经严格处理的废水直接排放，使得它们周围的土壤容易富集高含量的有毒重金属。企业排放的烟尘、废气中也含有重金属，并最终通过自然沉降和雨淋沉降进入土壤。矿业和工业固体废弃物在堆放或处理过程中，由于日晒、雨淋、水洗等，重金属极易移动，以辐射状、漏斗状向周围土壤扩散，固体废弃物也可以通过风的传播而使污染范围扩大。

2、交通运输对土壤重金属污染的影响

道路两侧土壤中的污染物主要来自汽车尾气排放及汽车轮胎磨损产生的大量含重金属的有害气体和粉尘的沉降，而污染元素则主要为 Pb、Cu、Zn 等元素。它们一般以道路为中心成条带状分布，强度因距离公路、铁路、城市以及交通量的大小有明显的差异。

5.3.3 定污染源的位置

首先，根据附件，使用 MATLAB 软件得到城区区域分布图（图 11）

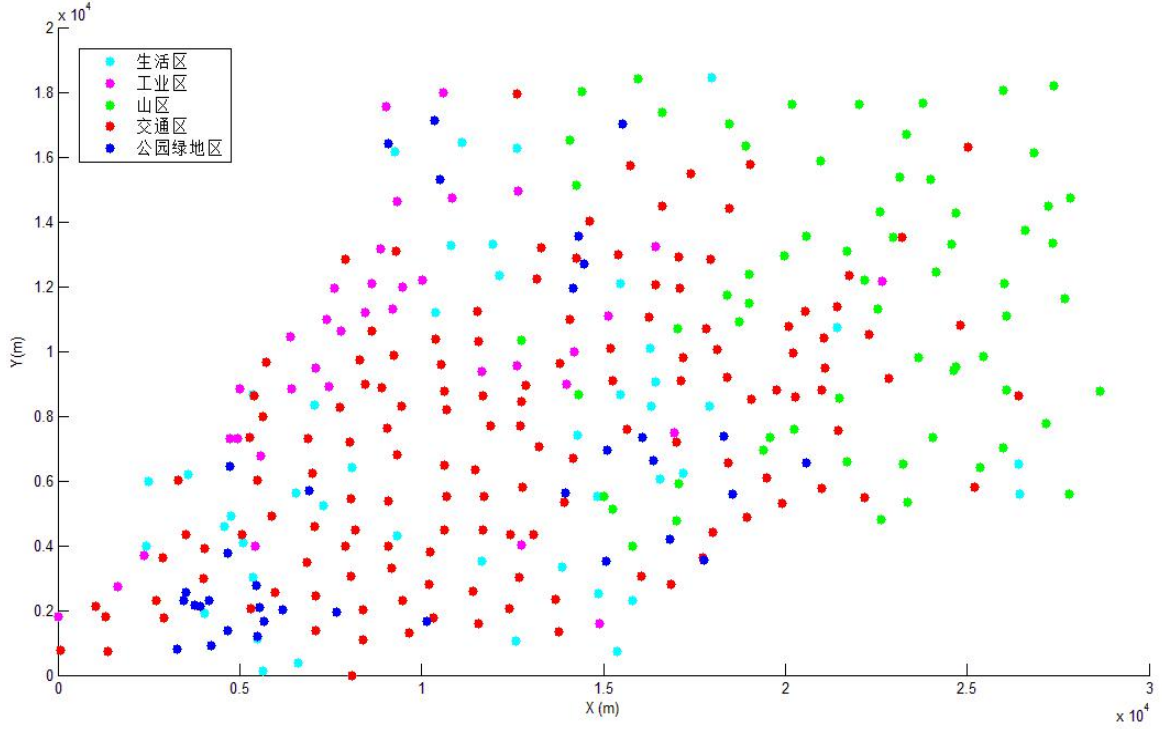


图 11 城区区域分布图

为简化模型，本文从 Cu 的污染分布着手，讨论污染源位置确定的模型，其他元素依次代入模型即可。

污染源 $(x_0, y_0, z_0) = (0, 0, 0)$

使用 INGLADA 方法进行初值计算^[4]

设观测点坐标为 (x_i, y_i, z_i)

由图 1—图 8 初步估计污染源多于 2 个

污染源到观测点的距离为 R

$$\text{对于污染源 } i, (x_0 - x_i)^2 + (y_0 - y_i)^2 + (z_0 - z_i)^2 = R_i^2 \quad (1)$$

$$\text{对于污染源 } j, (x_0 - x_j)^2 + (y_0 - y_j)^2 + (z_0 - z_j)^2 = R_j^2 \quad (2)$$

将 (1) (2) 式展开后相减得到线性方程：

$$(x_i - x_j)x_0 + (y_i - y_j)y_0 + (z_i - z_j)z_0 = \frac{1}{2}(r_i^2 - r_j^2 + R_j^2 - R_i^2) \quad (3)$$

其中 $r_i^2 = x_i^2 + y_i^2 + z_i^2$

由于土壤为均匀介质，传播速度恒定，得等式

$$\frac{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2}{(C_i - C_0)} = \frac{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2}{(C_j - C_0)} \quad (4)$$

将上述表格的数据每三个一组代入 (3) 式

得到一系列的 (x_0, y_0, z_0) ，从而反演得源位置。

由附件所得 320 组数据，每四个一组，共分成 80 个组，代入（4）式中，得到 80 组 (x_0, y_0, z_0) 污染源位置，进行统计后， x_0 的波动范围为 0—2454， y_0 的波动范围为 987

—1570，从而确定平面污染源的区域范围为 $\begin{cases} 0 \leq x_0 \leq 2454 \\ 987 \leq y_0 \leq 1570 \end{cases}$ 。

由题目中的采样点是在间距 1 公里左右的网格子区域内， $|z_i - z_j|$ 远小于 $|x_i - x_j|$ 和 $|y_i - y_j|$ ，（3）中的系数矩阵会有奇异性， z_0 解发散，这样更加体现此模型适宜这个问题。

1.2 用 Geiger 方法对初值进行迭代校正

基于 Geiger 法的基本思想，建立空间直角坐标系：

$$\Delta C + \frac{\partial C_k}{\partial x} \Delta x + \frac{\partial C_k}{\partial y} \Delta y + \frac{\partial C_k}{\partial z} \Delta z = r_k \quad (5)$$

式中 $\Delta x, \Delta y, \Delta z$ 分别为污染源坐标校正量， ΔC 为浓度校正量， $\frac{\partial C_k}{\partial x}, \frac{\partial C_k}{\partial y}, \frac{\partial C_k}{\partial z}$ 为空间偏导数， r_k 为测定浓度的偏差，可写成矢量形式：

$$A \Delta \vec{x} = \vec{r}, \quad (\text{其中 } \Delta \vec{x} = (\Delta C, \Delta x, \Delta y, \Delta z), \quad \vec{r} = (r_1, r_2 \cdots r_m)) \quad (6)$$

（6）式中系数矩阵 A 为 $m \times 4$ 矩阵：

$$A = \begin{pmatrix} 1 & \frac{\partial C_1}{\partial x} & \frac{\partial C_1}{\partial y} & \frac{\partial C_1}{\partial z} \\ 1 & \frac{\partial C_2}{\partial x} & \frac{\partial C_2}{\partial y} & \frac{\partial C_2}{\partial z} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{\partial C_m}{\partial x} & \frac{\partial C_m}{\partial y} & \frac{\partial C_m}{\partial z} \end{pmatrix} \quad (7)$$

将 Inglada 方法计算的初值代入（5）式求解，再用求解结果校正 r ，如此反复迭代校正，直至满足一下两个条件。

① $\Delta \vec{x} < \delta$ ， $\delta > 0$ 为任一小量

② 第 n 次迭代 $|\Delta x_n - \Delta x_{n+1}| < \varepsilon$ ， $\varepsilon > 0$ 为任一小量

（6）的求解方法有很多，比如牛顿-高斯算法^[5]，但对于此问题过于复杂，又比如为提高稳定性，可采用中心化、定标化、阻尼最小二乘法，但对于本题求解时，注意到 (x_0, y_0, z_0) 遍历的数据比较多，矩阵较大，所以本文采用共轭梯度法求解。

为标准评价土壤环境的污染程度，依据国家环保局颁布的土壤环境质量标准，土壤中各元素环境质量评价分级含量范围采用表 8 分级值。^[6]

表 8 土壤金属污染评价标准（标准对比法）

评价标准	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn
一类	≤15	≤0.2	≤90	≤35	≤0.15	≤40	≤35	≤100
二类	15-25	0.2-0.6	90-250	35-100	0.15-1.0	40-60	35-350	100-300
三类	25-40	0.6-1.0	250-300	100-400	1.0-1.5	60-200	350-500	300-500
超三类	>40	>1.0	>300	>400	>1.5	>200	>500	>500

5.4 问题四

5.4.1 分析优缺点

模型三中，我们大胆引进地震源的位置确定，采用的 Geiger 定位法是目前比较准确先进的定位方法。但是计算比较复杂。

5.4.2 收集信息

基于以上问题的分析，除了题目给出的四个变量：样方横坐标，样方纵坐标，海拔和采样点的浓度值，我们将条件理想化，建立了以上三个模型，为了更好的研究城市地质环境的演变模式，还要丰富很多信息和相关专业信息。我们知道，更加全面客观的评价某地区的土壤污染程度，还需要掌握该区域土壤特质即土壤的物理结构和化学结构，包括土壤由固相、液相、气相和 PH 值以及重金属元素的毒理性。

在模型二中我们采用主成分分析法对主要原因的分析，还应该收集该区域的地形地貌，每个区域的地理分布状况，土壤特性以及气候影响。当然也不能忽视每个区域日常活动状况，其中包括动物植物微生物。

在确定污染源的位置时，除了模型三所考虑的因素以外，重金属元素与化合物的相关性也应该纳入模型的计算，此外还应该考虑污染物与时间的关系，随着时间的推移，污染源的位置将随之变化，这样，静态模型就变为动态模型。

5.4.3 建立模型

我们会基于上述城市信息，采用城市的地质环境的数据，运用时间序列分析中的自回归方法建立城市地质综合环境质量演变差分趋势预测模型。

当预测对象随着时间呈现某种水平、上升或下降的趋势，并且无明显的季节波动，又能找到一条合适的曲线反映这种变化趋势时，就可用时间 t 为自变量，时序数值 y 为因变量，建立趋势模型： $y = f(t)$ ；

正确识别并选择趋势模型是首要工作，基本方法有两种：图形识别法、差分计算^[7]。在这里我们使用差分法，所谓差分，是变量的微小变化，根据历史数据计算差分把数据修匀，将非平稳时间序列转换为平稳序列，通过该平稳序列的表现，来发现该套用那个模型。

假设时间序列为 $y_t (t = 1, 2, 3, \dots, n)$

一阶向后差分为： $y_t' = y_t - y_{t-1}$

二阶向后差分为： $y_t'' = y_t' - y_{t-1}'$

三阶向后差分为： $y_t''' = y_t'' - y_{t-1}''$

多阶以此类推

再根据差分特性确定使用的模型。表（9）

表 9 差分特性对应的模型

差分特性	使用模型
一阶差分相等或大致相等	一次线性模型
二阶差分相等或大致相等	二次线性模型
三阶差分相等或大致相等	三次线性模型
一阶差分比率相等或大致相等	指数曲线模型
一阶差分的一阶比率相等或大致相等	修正指数曲线模型

其中，多次线性模型一般形式为：

$y = a_0 + a_1t + a_2t^2 + \dots + a_nt^n$ ， y 预测目标变量， t 为时间， b 为待定系数。当 $n=1$ 时，为线性模型； $n=2$ ，为二次抛物线模型； $n=3$ ，为三次多项式趋势模型。

估计标准误差 $SE: \delta = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p}}$ ，其中， n 为时间序列的长度， p 为限制条件，

即利用最小二乘法时的所列方程数，可以简单理解为直线模型为 2，抛物线模型为 3，三次曲线为 4，最高次数加 1。

指数曲线预测模型一般形式为：

$\hat{y} = ae^{bt}$ ($a > 0$)， a 为时间序列的初期水平，必须大于 0。 b 为时间序列的平均发展速度，可以为正，负值。

修正指数曲线预测模型后，得 $\hat{y}_t = k + ae^{bt}$ ，其中 t 为时间， $k > 0, 0 < b < 1$ ，一阶差分的环比为常数 b ， $\frac{\nabla_t}{\nabla_{t-1}} = \frac{\hat{y}_t - \hat{y}_{t-1}}{\hat{y}_{t-1} - \hat{y}_{t-2}} = b$ ，待定系数 a, b, k 的求法采用“三段和值法”进行。

六、模型评价与推广

6.1 模型的优点

在问题一中，我们采用目前学术界最常用的单因子指数法可以判断出环境中的主要污染因子，是综合评价指数的基础。为了更加全面合理地进行污染程度分析，我们采用尼梅罗综合指数法进行综合评价。它不仅考虑了各重金属对土壤作用的平均水平，更重要的是突出了较高浓度污染物对土壤环境质量的影响，因此，高浓度污染物对评价结果具有较大影响。这样突出了污染指数最大的污染物对环境质量的影响和作用，而且此方法操作简单，容易实现，给问题分析和结论提供了很好的依据。

模型二中我们利用主成分分析法 进行精确计算，得出主成分。

模型三中我们大胆引进地震源的位置确定，采用的 Geiger 定位法是目前比较准确先进的定位方法。

模型四中我们发散思维，综合考虑了不同侧面，对问题有一个全面客观的预测。

6.2 模型的缺点

单因子指数法于尼梅罗指数法只能反映污染的程度而难于反映污染的质变特征^[8]。

模型二中的主成分分析法沿用了模型一中的结论，这样对模型一的依赖性就增强了，这会对模型造成使用的局限。

参考文献

- [1] 于小俸 唐崑 胡玉, 壶瓶山自然保护区土壤重金属含量的空间分布及污染评价, 中国农学通报, 26(11): 358-361, 2010。
- [2] 闫欣荣, 修正的内梅罗指数法及其在城市地下饮用水源地水质评价中的应用, 地下水, 32(1): 6-7, 2010。
- [3] 邵学新 吴 明 蒋科毅, 土壤重金属污染来源及其解析研究进展, [http://journal.9med.net/qikan/article.php?id=296504\(2011-9-10\)](http://journal.9med.net/qikan/article.php?id=296504(2011-9-10))
- [4] 牟磊育 赵仲和 张伟, 用 INGLADA 与 GEIGER 方法实现近震精定位, 中国地震, 22(3): 294-302, 2006。
- [5] 林峰 李庶林 薛云亮 徐宏斌, 基于不同初值的微震源定位方法, 岩石力学与工程学报, 29(5): 996-1002, 2010。
- [6] 杨贤智, 李景铨, 廖延梅, 环境管理学, 北京: 高等教育出版社, 1990。
- [7] 姜启源 谢金星 叶俊, 数学模型, 北京: 高等教育出版社, 2003。
- [8] 徐燕 李淑芹 郭书海 李凤梅 刘婉婷, 土壤重金属污染评价方法的比较, 安徽农业科学, 36(11): 4615-4617, 2008。