辐射环境监测优化布点的特征分析方法

史东生，周春林，弟宇鸣，黎素芬
(第二炮兵工程学院 102教研室，陕西　西安　710025)

摘　要：介绍了特征分析方法的基本原理及其 MATLAB语言实现过程，并将特征分析法、传统布点法，以及 BP人工神经网络法对同一地区辐射环境监测优化布点的核素含量均值作了比较，结果表明，用特征分析法优选出的采样点位，其核素含量均值对传统布点法的相对偏差均为负值，且 < 10%，即便与 BP法相比，也不失其优越性，表明用特征分析法监测一个地区环境辐射平均水平的优化布点是十分理想的。

关键词：辐射；环境监测；布点；特征分析方法；MATLAB语言
中图分类号：X830.1　　文献标识码：C　　文章编号：1006－2009(2006)04－0036－03

Optimal Number of Sites of Radiological Environmental Monitoring Based on Characteristic Analysis

SHI Dong-sheng, ZHOU Chun-lin, DI Yu-ming, LI Su-fen
(*102 Research Group, The Second Artillery Engineering Institute, Xi'an, Shanxi 710025, China*)

Abstract: This paper introduced a new way-characteristic analysis with its basic principle and the steps of building its MATLAB program. Moreover, the characteristic analysis was compared with others measures, such as traditional sites and BP NN in the field of environmental radiation monitoring. The results of application examples suggested that it was feasible to optimize number of sites for monitoring the average level of environmental radiation in one area by characteristic analysis.

Key words: Radiological; Environment monitoring; Number of sites; Characteristic analysis method; MATLAB Program

在辐射环境质量评价中，对现场监测与采样点设置的优化是一个重要的技术问题[1]。优化布设采样点的目的是尽可能合理、准确、完整地反映区域辐射水平的空间分布和演变规律。常用的优化布点方法有系统聚类法、物元分析法[2]和模糊优化法[3]等，这些方法都未能很好地解决多因素指标之间非线性权值分配问题。近年发展起来的人工神经网络方法能够向不完整、不精确并具有噪声的样本学习，具有非常强的容错能力，能够从有限的信息中得出最优解，显示了人工神经网络方法具有较强的优越性[4]。现根据特征分析模型[5]提出一种新的方法即特征分析法，该方法可以用于监测一个地区环境辐射平均水平的优化布点，与人工神经网络方法（简称 BP法）相比，具有异曲同工之效。

1　特征分析方法原理

1.1　特征分析模型

设有样本 n 个，变量 m 个，记为 x_j，变量 x_j 在第 i 个样本上的取值记为 x_{ij} ($i=1, 2, \cdots, n$; $j=1, 2, \cdots, m$)。

样本的联系度 $y=(y_1, y_2, \Lambda, y_n)'$ 定义为诸变量的线性组合：

$$y=a_1x_1+a_2x_2+\Lambda+a_mx_m \qquad (1)$$

式中：a_1, a_2, Λ, a_m ——m 个变量相应的权值；

y——衡量样本之间相似性程度的一种度量，称之为样本联系度。

由(1)式：

收稿日期：2005－12－06；修订日期：2006－04－24
作者简介：史东生（1978—），男，河北保定人，在读硕士研究生，从事辐射防护与环境监测研究。

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \tag{2}$$

为求各变量权值 $(j=1, 2, \cdots, m)$, 应使已知样本联系度最大, 定义下式:

$$Q = \max \sum_{i=1}^{n} y_i^2 \tag{3}$$

由此, (2)式和(3)式的矩阵形式应为:

$$\begin{cases} \vec{y} = (y_1 \quad y_2 \quad \cdots \quad y_n)' \\ x = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \\ \vec{a} = (a_1 \quad a_2 \quad \cdots \quad a_m)' \end{cases} \tag{4}$$

于是, (2)式可写成:

$$\vec{y} = x\vec{a} \tag{5}$$

由(3)式求解方程(5)中变量权值向量 $\vec{a}$ 的问题归结为解方程组:

$$\begin{cases} \max \|\vec{y}^2\| \\ \|\vec{a}^2\| = 1, \ 0 \leqslant a_j \leqslant 1, \ j = 1 \cdots m \end{cases} \tag{6}$$

由拉格朗日乘因子法, 满足(6)式的解 $\vec{a}$ 应满足:

$$\begin{cases} x'x\vec{a} = \lambda\vec{a} \\ \vec{a}\vec{a} = 1 \end{cases} \tag{7}$$

当时 $\lambda = \max(\|x\vec{a}^2\|)$ 才使已知样本联系度最大, 此时, λ 为矩阵 $x'x$ 的最大特征根, $\vec{a}$ 则为相应的特征向量。

1.2 联系度

依据上述变量权值向量 $\vec{a}$ 的计算方法可以计算出各变量的权值, 根据线性方程组

$$y_i = \sum_{j=1}^{m} a_j x_{ij} \quad (i = 1, 2, \cdots n; \ j = 1, 2, \cdots m) \tag{8}$$

可以计算出各样品的联系度。

2 特征分析法的应用及 MATLAB 语言实现

2.1 原始数据的“二进制”赋值

采用文献[6]的数据, 24个采样点的 ^{238}U、^{226}Ra、^{232}Th 和 ^{40}K 4种放射性核素监测值见表 1。

表 1中的 u、v、w 分别表示“最佳理想点”、“最次理想点”和“数学期望点”3个参考点, 其中:

$$u_j = \min x_{ij}; \quad v_j = \max x_{ij}; \quad w_j = \sum_{i=1}^{n} x_{ij} / n \tag{8}$$

表 1　某地区辐射环境土壤样品放射性核素含量

Bq/kg

编号	^{238}U	^{226}Ra	^{232}Th	^{40}K
01	27.60	41.52	59.24	601.71
02	19.62	33.62	47.52	516.97
03	20.56	40.84	60.51	690.61
04	27.02	43.77	55.06	622.17
05	32.18	44.67	60.19	523.05
06	29.53	43.38	69.83	587.18
07	23.81	36.18	50.55	564.67
08	28.55	47.28	60.75	614.18
09	30.48	38.85	54.93	563.04
10	38.49	47.00	53.78	478.61
11	44.25	68.79	68.75	564.46
12	17.67	38.95	62.93	606.15
13	28.96	40.10	59.10	439.50
14	24.80	30.00	60.75	619.50
15	27.00	36.24	52.84	538.40
16	20.36	25.33	39.21	505.55
17	22.73	30.52	43.38	517.69
18	18.97	26.94	32.90	416.71
19	28.63	51.75	59.00	539.33
20	27.58	37.69	54.77	584.08
21	48.63	71.83	76.38	575.29
22	21.56	31.16	44.24	555.76
23	47.36	48.91	64.18	615.50
24	22.48	34.74	60.96	567.56
u	17.67	25.33	32.90	416.71
v	48.63	71.83	76.38	690.61
w	28.29	41.26	56.32	558.65

式中 $i = 1, 2, \cdots n$, n 为采样点数目; $j = 1, 2, \cdots m$, m 为监测指标数目。“二进制”赋值规则是对某个样本和参考点:

(1) 如果某核素监测值接近相应的 u 点值, 则其赋值为(1 0 0);

(2) 同理, 某核素监测值接近相应的 w 点值, 则其赋值为(1 1 0);

(3) 监测值接近 v 点值, 则其赋值为(1 1 1)。

以 01号采样点为例, 它的“二进制”赋值为(1 1 0 1 1 0 1 1 0 1 1 0), 这样表 1中 24个采样点经过变换之后, 即可得到它们在各变量上的赋值。

2.2 MATLAB 语言实现

MATLAB 程序如下:

X = [1 1 0 1 1 0 1 1 0 1 1 0, 1 0 0 1 1 0 1 1 0 1 1 0; 1 0 0 1 1 0 1 1 0 1 1 1; 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 1 1 1 0, 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 1 1 1 0 1 1 0 1 0 0, 1 1 1 1 1 1 1 1 1 1 1 0, 1 0 0 1 1 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 0 1 0 0, 1 1 0 1 0 0 1 1 0 1 1 0, 1 1 0 1 1 0 1 1 0 1 1 0; 1 0 0 1 0 0 1 0 0 1 1 0, 1 0 0 1 0 0 1 0 0 1 1 0, 1 0 0 1 0 0 1 0 0 1 0 0, 1 1 0 1 1 0 1 1 0 1 1 0; 1 1 0 1 1 0 1 1 0 1 1 0, 1 1 1 1 1 1 1 1 1 1 1 0, 1 0 0 1 0 0 1 0 0 1 1 0, 1 1 1 1 1 0 1 1 0 1 1 0, 1 0 0 1 1 0 1 1 0 1 1 0]; % 24个采样点经过变换之后的“二进制”赋值矩阵。

```
X = x'* x;
[A D] = eig(X); % 求矩阵 X 的特征向量矩阵 A 与特征值矩阵 D
[n m] = size(x);
z = -10 000;
for i = 1:m
    if z < D(i, i)
        z = D(i, i);
        j = i;
    end
end
a = A(:, j); % 求对应 X 矩阵最大特征值的特征向量
y = x* a; % 求各样本的联系度
y = ';
t = 1:n;
plot(t, y, t, y, 'o'); % 绘图显示运行结果
xlabel('样本编号');
ylabel('样本联系度');
grid
```

运行程序后, 结果加上标注显示, 见图 1。

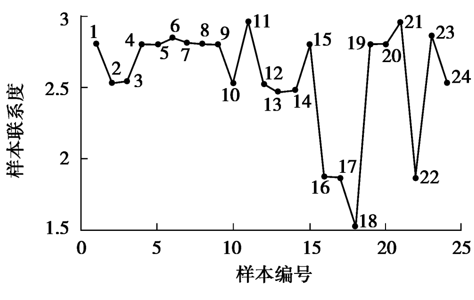

图 1 样本联系度曲线

3 结果与分析

由图 1可清楚地看到 24个样本联系度的大小, 根据各样本的联系度值将其大致分为 {18}、{16, 17, 22}、{13, 14}、{02, 03, 10, 12, 24}、{01, 04, 05, 07, 08, 09, 15, 19, 20}、{06, 23}和 {11, 21}共 7组, 这样就可结合实际工作确定优化点位, 现选取 18、16、13、02、01、06、11号样本作为优化点, 从理论上讲也是合理的。用特征分析法、传统布点法和 BP法对同一地区辐射环境监测优化布点的核素含量均值比较见表 2。

表 2 特征分析法、传统布点法和 BP法的各核素含量均值比较

方法	点位数	^{238}U	^{226}Ra	^{232}Th	^{40}K
传统布点法 a/(Bq·kg^{-1})	24	28.29	41.26	56.32	558.65
BP法[4] a/(Bq·kg^{-1})	9	28.56	44.89	57.30	630.06
相对偏差/%①		0.95	8.8	1.7	12.8
特征分析法 a/(Bq·kg^{-1})	7	27.04	39.95	53.79	518.87
相对偏差/%②		-4.4	-3.2	-4.5	-7.1

①以传统布点法为参考值的 BP法相对偏差; ②以传统布点法为参考值的特征分析法相对偏差。

由表 2可见, 用特征分析法优选出的采样点位, 其核素含量均值对传统布点法的相对偏差均为负值, 且 < 10%。即便与 BP法相比, 也不失其优越性。

[参考文献]

[1] 刘明, 吕魁, 张斌. 田湾核电站原野外 γ辐射剂量率调查[J]. 环境监测管理与技术, 2003, 15(6): 17-20.

[2] 于永斌, 朱慧君. 大气监测优化布点的物元分析法[J]. 环境保护, 1997, (6): 25-27.

[3] 刘全有. 环境监测点的模糊优化设置问题探讨[J]. 环境科学进展, 1996, 17(5): 41-43.

[4] 周春林. 辐射环境监测优化布点的 BP人工神经网络方法[J]. 辐射防护, 2005, 25(3): 163-167.

[5] BOTBOL JM. An Application of Characteristic Analysis to Mineral Exploration[J]. Proc gth Int, 1971.

[6] 周春林. 二炮阵地环境放射性水平调查与评价关键技术研究[D]. 西安: 第二炮兵工程学院, 2004.