

# 天津市塘沽区地面沉降监测网络优化

董国凤<sup>1</sup> 张 蕾<sup>1</sup> 赵 全<sup>1</sup> 赵新华<sup>1</sup>

(1 天津大学环境科学与工程学院, 天津市卫津路 92 号, 300072)

**摘 要:**以天津市塘沽区为研究区域, 利用空间聚类与数据序列特征分析相结合的方法, 对原有 86 个监测点的地面沉降监测网络进行优化, 使优化兼顾了空间均匀分布和最小化时间冗余。给出了监测点数目, 仅保留其中大约 25% 的监测点进入优化后的监测网络。

**关键词:**地面沉降; 监测网络优化; 塘沽区

**中图分类号:** P258

长期以来, 由于超采地下水造成的破坏性的地面沉降已经成为制约地方经济发展全局的复杂环境地质问题<sup>[1-4]</sup>。政府有关部门在许多地面沉降危害严重的地区建立了地面沉降监测网络, 然而维持规模庞大的监测网络不可避免地消耗大量的人力和物力; 另一方面, 随着控制地面沉降(简称控沉)工作的深入和可持续发展观念的建立, 许多一度经历破坏性地面沉降的地区逐渐进入微量沉降(年沉降量 1~2 cm)阶段<sup>[5]</sup>, 年沉降量趋于稳定。通过优化地面沉降监测网络, 可以降低维持费用, 提高控沉工作的经济效益, 使有限的资源和投入得到更加合理的配置。

天津市塘沽区东临渤海湾, 西距市中心 45 km, 建成区 188 km<sup>2</sup>。执行控沉计划以来, 区内的地面沉降平均速率由急剧发展阶段(80~150 mm/a)降至 8~32 mm/a 的稳定水平。本文利用空间聚类与数据序列特征分析相结合的方法, 对区内及周边具有完整时间序列的 86 个监测点的地面沉降监测网络进行了优化研究。在兼顾空间均匀分布和最小化时间冗余的目标下, 尽可能减少监测点的数目并降低平时的维护费用。

## 1 监测网络优化

### 1.1 空间聚类分析

利用 Matlab 软件提供的  $K$ -均值函数<sup>[6]</sup>对所有 86 个地面沉降监测点的地理坐标进行聚类分

析。 $K$ -均值聚类方法将观测值视为空间中的点, 在给定分组数量的前提下采用欧氏距离求算各个元素所在分组。在随机选定初始分组后, 通过迭代不断修正各个元素的分组情况, 使得每组内的元素彼此尽量接近, 而不同组间元素尽可能彼此远离。

对于每一个元素  $i$ , 其分组的质量优劣以 Silhouette 值(简称为 Silh, 其值设为  $q_{\text{silh}}$ )表示<sup>[6]</sup>, 定义如下:

$$q_{\text{silh}}(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (1)$$

式中,  $a(i)$  是某元素到同组元素的平均距离;  $b(i)$  是某元素到最近的组中所有元素的平均距离。 $q_{\text{silh}}(i)$  用以衡量  $b(i)$ 、 $a(i)$  之间的标准差, 其值介于 -1 和 1 之间。 $q_{\text{silh}}(i)$  接近 1, 那么样本  $i$  距离同组内的各元素比距离其他邻近的组近, 所以分类良好; 反之, 如果接近 -1, 则是被错分的。Kaufman 等认为, 值超过 0.5 的  $q_{\text{silh}}$  就是好的分类结果, 0.2 以下缺少实质聚类结构。 $S$  定义为  $q_{\text{silh}}(i)$  的均值, 用来表示聚类分析的整体分散程度:

$$S = \sum_{i=1}^n q_{\text{silh}} / n, \quad n = 86 \quad (2)$$

同样,  $S$  越接近 1, 则表示该聚类方案的整体分散程度越强, 元素的分散情况更趋于合理。

设定分组数量是  $K$ -均值聚类分析的前提, 将对聚类分析的结果产生决定性的影响。由于分组起始点是随机选定的, 每次运算的结果可能由于

落入局部最优而不同。通常,这种情况是不可避免的,除非出现幸运起始点或穷尽所有可能的起始情况。聚类分析的分散水平和稳定性水平随着分组数量的改变而不断改变。所以,往往通过选择多次重复计算试探最优解,并决定合适的分组数量。

通过5次和10次重复试探计算的实验结果可以得出,随着分组数目  $k$  由3增加到20,  $S$  均值有增大的趋势,表示聚类分析的分散度增加;当分组数超过7时,5次和10次计算的  $S$  极大值出现不同,表示局部最优出现的可能性增加,且全局最优越来越不易确定。经过对  $k=3, \dots, 20$  各种分类模式结果的比较,选择了  $k=7$  作为最终聚类分组的依据,同时兼顾了稳定性和分散水平的要求。聚类分析  $k=7$  时的 Silh 分布见图1,地面沉降监测点聚类分析结果见图2。

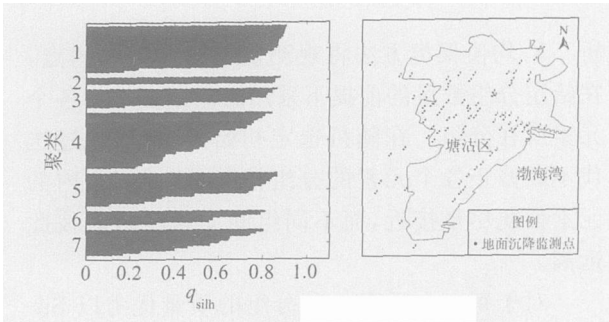


图1  $k$  为7时的 Silh 分布

图2 地面沉降监测点聚类分析结果

Fig. 1 Distribution of Silh when  $k$  is 7

Fig. 2 Map of Results of Cluster Analysis of Land Subsidence Monitoring Points

### 1.2 数据序列分析

为了减少监测点以降低维护费用,使优化网络具备原有地面沉降监测网络时空分布的监测能力,依据塘沽区的地域面积及其目前的沉降特点,在聚类分析的基础上从每一分组中选取约25%的监测点进入优化后的监测网络,见表1。

表1 优化前后监测点数量变化

Tab. 1 Variation of Monitoring Points Before/After Optimization

所属分组	1	2	3	4	5	6	7	合计
化前数量	20	3	8	24	14	6	11	86 100%
化后数量	5	1	2	6	3	2	3	22 25.6%

利用标准偏差表征各个监测点年沉降量历史数据序列(1986年~2003年)的离散程度,从每个聚类分组中优先选择标准偏差大的监测点进入优化后的监测网络,直到每个分组中的监测点数量达到优化监测点数量限制。最终的监测网络优化结果如图3。



图3 地面沉降监测网络优化结果

Fig. 3 Map of Optimization Results of Land Subsidence Monitoring Network

## 2 优化结果检验

为了检验优化后的22点监测网络与原有86点监测网络的监测效果,选取塘沽区2000年和2002年度的年地面沉降量观测数据,利用 ArcGIS 8.2 软件进行二阶反距离权重插值<sup>[7]</sup>,以两种监测网络插值结果得出的各个地面单元(插值空间单元)的相对偏差  $q$  以及整体平均相对偏差  $P$  为评价指标。定义如下:

$$q_i = | (S_{o_i} - S_{N_i}) / S_{o_i} | \quad (3)$$

$$P = ( \sum_{i=1}^m q_i ) / m \quad (4)$$

式中,  $S_{o_i}$  代表86点监测网络插值后任一地面单元的年地面沉降量;  $S_{N_i}$  代表22点监测网络插值后任一地面单元的年地面沉降量;  $q_i$  代表任一地面单元优化前后地面沉降插值结果的相对偏差;  $m$  为研究区内所有地面单元的总数;  $P$  为整体平均相对偏差。 $P$  值实际上代表了优化后的网络与原有网络监测年地面沉降量插值结果的差别,这些差别是由于精简了部分(约75%)监测点造成的。优化网络与原有网络年沉降量插值结果在2000年和2002年度平均相对偏差分别为18.1%和17.8%。2000年和2002年度  $q$  值分布见表3。

表2 2000年及2002年  $q$  值分布

Tab. 2 Distribution of  $q$  Value in 2000 and 2002

$q$ 值域分布	$q$ 值域所属地面单元占总监测区域比例/%	
	2000年	2002年
$q \leq 5\%$	23.3	20.5
$q \leq 10\%$	38.5	39.9
$q \leq 20\%$	61.2	69.5
$q \leq 30\%$	77.7	82.6
$q > 30\%$	22.3	17.4

可见大多数地面单元插值数据的相对偏差控制在 30% 以内,保证了优化后监测网络的可靠性。年沉降量等值线图分别见图 4。

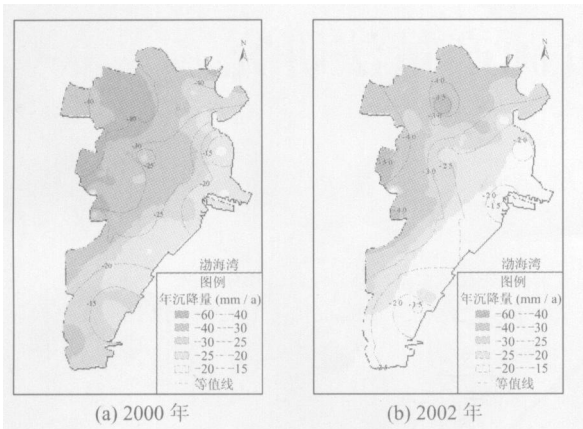


图 4 塘沽区地面沉降等值线图

Fig. 4 Contour Map of Land Subsidence at Tanggu

图 4 底图中色阶所表示的是 86 点监测网络地面沉降观测值插值结果,图中等值线是由 22 点优化网络地面沉降观测值插值结果绘制的。可见,优化后的 22 点监测网络在给定监测点数量的前提下兼顾了空间均匀分布和最小化时间冗余的目标,继承了原有 86 点网络对地面沉降时空分布的监测能力。

### 3 结 语

为了兼顾监测网络空间均匀分布和最小化时间冗余的目标,本研究将空间聚类与数据序列特征分析相结合的方法首次应用在天津市塘沽区

86 个监测点的地面沉降监测网络的优化,根据塘沽区的地域面积及其地面沉降的特征,保留了其中 22 个监测点进入优化后的监测网络。优化后的监测网络可在降低维护费用的前提下,继承原有 86 点网络对地面沉降时空分布的监测能力,并可以显著提高微量沉降阶段控沉工作的经济效率。

### 参 考 文 献

- [1] 段永侯,王家兵,王亚斌,等. 天津市地下水资源与可持续利用[J]. 水文地质工程地质, 2004(3): 29-39
- [2] 吴铁均,崔小东,牛修俊,等. 天津市地面沉降研究及综合治理[J]. 水文地质工程地质, 1998(5): 17-20
- [3] 杨建图. 天津市地面沉降防治对策战略研究[J]. 工程地质学报, 2005(13): 4-7
- [4] 刘毅. 地面沉降研究的新进展与面临的新问题[J]. 地质前缘, 2001, 8(4): 273-277
- [5] 刘杜娟. 中国沿海地区地面沉降的危害及防治对策——以天津市为例[J]. 海洋地质动态, 2004, 20(1): 1-7
- [6] Kaufman L, Rousseeuw P.J. Finding Groups in Data: an Introduction to Cluster Analysis [M]. New York: John Wiley & Sons, 1990
- [7] Chang Kang tsung. 地理信息系统导论[M] //陈健飞,译. 北京: 科学出版社, 2003

第一作者简介:董国凤, 博士生, 研究方向为环境系统优化与数学模拟。

E-mail: tjszyc@eyou.com

## Land Subsidence Monitoring Network Optimization at Tanggu in Tianjin

DONG Guofeng<sup>1</sup> ZHANG Lei<sup>1</sup> ZHAO Quan<sup>1</sup> ZHAO Xinhua<sup>1</sup>

(1 School of Environmental Science and Engineering, Tianjin University, 92 Weijin Road, Tianjin 300072, China)

**Abstract:** A method combining spatial cluster with data statistical analysis was employed to optimize a land subsidence monitoring network consisted of 86 monitoring points at Tanggu in Tianjin. So the optimized network met with the goal of both maximizing spatial accuracy and minimizing temporal redundancy. Giving the number of monitoring points, an optimized network containing a selection of about 25% monitoring points from the original network was built.

**Key words:** land subsidence; monitoring network optimization; Tanggu District

About the first author: DONG Guofeng, Ph. D candidate, majors in environmental system optimal and mathematics mode.

E-mail: tjszyc@eyou.com