

湖泊水质参数空间最优估计精度分析

刘瑞民¹, 王学军^{1*}, 王翠红² (1. 北京大学城市与环境学系, 北京 100871, E-mail: rm-liu@263.net; 2. 济南市舜耕中学, 济南 250014)

摘要:在湖泊研究中, 往往要了解污染的空间分布状况, 其核心是空间最优估计问题. 本文在阐述了克立格方差特点的基础上, 以太湖为例, 对湖泊水质参数空间最优估计的精度问题进行了分析. 结果显示, 估计精度受变差函数、各采样点的空间配置关系等因素的影响, 空间结构直接决定着估计精度的大小, 而与各信息采样点的值的大小无关.

关键词:水质参数; 湖泊; 空间估计精度; 克立格方差

中图分类号: X524 **文献标识码:** A **文章编号:** 0250-3301(2001)05-04-0091

The Accuracy Analysis on Kriging Estimation of the Water Quality Parameters of Lakes

Liu Ruimin¹, Wang Xuejun^{1*}, Wang Cuihong² (1. Depart. of Urban and Environmental Sciences, Peking Uni., Beijing 100871, China E-mail: rm-liu@263.net; 2. Jinan Shungeng Middle School, Jinan Shandong 250014, China)

Abstract: The spatial distribution of pollutants in lakes is required in the study of lakes. The key issue is the spatial optimal estimation. This paper elaborated the characteristics of Kriging variance and analyzed systematically the accuracy of Kriging estimation on water quality parameters of lakes in the light of a practical example of Taihu Lake. The accuracy of estimation was determined by many factors such as variogram, distribution of sampling sites and so on, and was not related to the values of sampling data.

Key words: water quality parameter; lake; accuracy of spatial estimation; Kriging variance

在湖泊研究中, 经常要遇到了解各水质参数在湖水中的空间分布情况, 其核心问题就是空间最优估计. 而估计精度的大小则关系到对水质参数空间分布的了解以及后续各种分析和研究的准确性. 因此有必要对湖泊水质参数空间最优估计的精度问题进行深入研究.

地统计学是近几十年兴起的一门处理空间数据的数学方法, 主要包括空间结构分析和空间估计(克立格法)²部分^[1]. 与传统估计方法相比, 克立格法是一种最优、无偏内插估计方法. 同时克立格法在湖泊水质参数不确定性和离散性研究中的应用, 提供了衡量估计精度的定量指标即克立格方差.

从分析克立格方差的特点出发, 讨论湖泊水质参数随机场空间最优估计的精度问题, 对

于完善湖泊水质参数空间变异性分析具有特殊的意义.

1 度量空间最优估计精度的定量指标及特征

地统计学中一个重要的概念是区域化变量, 将随空间位置 x 而变化的湖泊水质参数看作是一种区域化变量 $Z(x)$. 这样就可以应用地统计学的理论与方法对湖泊水质参数空间最优估计的精度问题尝试进行分析.

1.1 空间最优估计精度的定量指标

基金项目: 国家自然科学基金资助项目(40041002); 教育部骨干教师资助计划项目.

作者简介: 刘瑞民(1975~), 男, 硕士, 主要从事环境信息系统方面的研究.

收稿日期: 2000-12-28

* 通讯联系人

一个好的估计方法不只是简单地给出一个待估点的估计值,而且还应给出一个确定估计值可靠性的数量指标,即估计精度,构成一个完整的空间估计体系.最优估计的克立格法与传统的估计方法不同,它能提供必要的估计精度——克立格方差,其数学表达式为^[2]:

$$\sigma_K^2 = \sum_{i=1}^n \lambda_i \gamma(h_{i0}) + \mu \tag{1}$$

式中, σ_K^2 表示克立格方差; λ_i 表示克立格权重系数; $\gamma(h_{i0})$ 表示各采样点 x_i 与待估点 x_0 之间的变差函数值; μ 是拉格朗日乘数.其中, λ_i 、 $\gamma(h_{i0})$ 均可由湖泊水质参数空间结构分析结果获得.

1.2 克立格方差的特征

从克立格方差的表达式可以看出,克立格方差的大小主要受采样点与待估点之间的变差函数值 $\gamma(h_{i0})$ 、各采样点的克立格权重系数 λ_i 以及拉格朗日乘数 μ 等因素的影响,而与各采样点值的大小无关.因而只要在数据构型和结构构型不变的条件下,便可以在待估点取得实测值之前根据变差函数值以及普通克立格方程组^[3]的求解结果,事先计算出湖水中待估点的估计精度.

2 空间最优估计的精度分析

2.1 不同空间估计方法之间的比较

为了表明在进行湖泊水质参数空间分布的估计时,克立格法确比其它传统空间估计方法有更强的有效性,笔者分别运用克立格法、线性内插法和距离平方反比法对太湖水域内随机选取的10个采样点的水质参数总悬浮物进行了估值计算(采样时间为1999-11-04,在太湖中均匀布置12行10列共75个采样点,由国家环保局太湖流域环境监测网中心站进行采样和测定,采样和样品的分析根据有关规范进行).其对比结果无论在估计误差均值(线性内插法为-1.05,距离平方反比法为-1.17,克立格法为-0.47)还是在估计误差方差(线性内插法为4.92,距离平方反比法为5.15,克立格法为2.22)上都明显表明克立格法是空间估计的最

优方法.对比结果见图1.

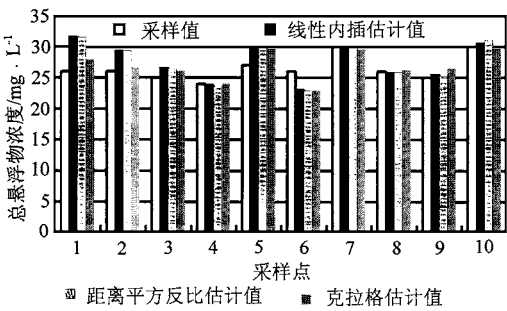


图1 克立格方法与传统估计方法结果比较
Fig.1 Comparison between estimation results by Kriging and traditional methods

2.2 空间最优估计的精度分析及结果

为了分析湖泊水质参数空间最优估计的估计精度,本文应用太湖的一些采样数据,采用克立格法对检验采样点进行空间最优估计,并对检验点的克立格方差及估计误差进行对比分析.

(1)克立格方差是指在特定数据构型和空间结构下,利用各种采样数据进行空间估计的估计方差的取值中心,即“期望值”,因而表现出与参估采样值大小无关并可事先进行精度预测的特点^[3~5],是衡量湖泊水质参数空间最优估计精度的指标.一般情况下,当待估点位于不同位置时,即已知样点与待估点之间的几何关系发生变化时,克立格方差值的大小发生变化.当待估点位于各采样点之间时,估计精度高一些,而当待估点在各采样点的一侧或采样点较少时,估计精度低一些.另外,当待估点距离采样点较远时,估计精度较低,待估点距采样点较近时估计精度较高(表1).

(2)区域化变量 $Z(x)$ 的空间变异程度影响估计精度^[5].克立格法中一个重要的概念是区域化变量,湖泊水质参数可以看作是一种区域化变量,它的变异性体现了空间结构的变异性.利用太湖中的采样结果,求得不同水质参数区域化变量空间变异性与估计效果,结果见表2.从表2可以明显地看出,不同的水质参数区域化变量空间变异程度存在着明显的差别,这种

差异导致各种水质参数的变差函数表达式 $\gamma(h)$ 不一样,最终影响克立格方差值的大小. 比较不同的水质参数区域化变量 $Z_{(x)}$ 空间估计精度与误差(表 3 和表 4),可以明显看出空间结构的变异性越大,其估计精度越小. (3)在湖泊水质研究中,由于受到湖岸线和

表 1 不同位置区域化变量估计精度对比¹⁾

Table 1 Kriging variance of regionalized variables at different locations						
变量 预测点号	总悬浮物(Z_1)		透明度(Z_2)		叶绿素 a(Z_3)	
	A_1	A_2	A_1	A_2	A_1	A_2
预测点相 x	17.766	23.342	17.766	23.342	17.766	23.342
对坐标/km y	23.306	61.183	23.306	61.183	23.306	61.183
预测值	3.110	3.429	3.746	3.881	-5.444	-3.548
σ_K^2	0.359	0.389	0.268	0.317	0.621	0.702

1) 各水质参数的监测值及预测值均为取对数后的值,下同.

表 2 区域化变量空间变异性与估计效果对比¹⁾

Table 2 Comparison between spatial variability of regionalized variables and estimated Kriging accuracy							
$Z_{(x)}$	D_V				$\overline{D_V}$	$\overline{\sigma_K^2}$	估计相对 误差/ %
	E-W	NE-SW	S-N	NW-SE			
Z_1	0.835	0.301	0.630	0.615	0.595	0.366	6.784
Z_2	0.603	0.281	0.453	0.500	0.459	0.275	0.614

1) D_V 为多元信息方向综合变异指标^[6]

表 3 区域化变量(Z_1)空间估计结果检验

Table 3 The tests in estimated Kriging accuracy of regionalized variables (Z_1)					
点号	实测值	估计值	克立格 方差	误差	相对误差/ %
1	3.178	3.147	0.359	-0.031	0.754
2	2.773	3.004	0.359	0.231	8.330
3	3.932	3.128	0.366	-0.803	20.422
4	3.296	3.333	0.376	0.037	1.123
5	3.434	3.321	0.369	-0.113	3.291

湖中岛等因素的影响,在靠近这些地方的待估点的信息点空间分布将发生变化,导致数据构型发生变化,从而影响到估计精度.对比太湖湖岸和湖中岛附近的待估点和湖中心的待估点(表 5 和表 6),可以明显地看出,在湖中心附近的估计精度高,而靠近湖岸和湖中岛的估计精度低. (4)在用克立格法对湖泊水质参数区域化变量进行估计时,不会因某个区域的采样样品丛聚在一起而过分增大这一区域采样样品总的克立格权系数,这就是所谓的减弱丛聚效应^[7].此外,克立格权重系数还具有屏蔽效应,即当块金常数为 0 或很小时,距离待估点最近

1 圈采样点的权系数明显地要比距待估点周围第 2 圈采样点的大一些,但这种屏蔽效应随着块金常数值增大而逐渐减小^[5,8].

表 4 区域化变量(Z_2)空间估计结果检验

Table 4 The tests in estimated Kriging accuracy of regionalized variables (Z_2)					
点号	实测值	估计值	克立格 方差	误差	相对误差/ %
1	3.912	3.897	0.268	-0.015	0.383
2	3.912	3.945	0.268	0.033	0.844
3	4.094	4.075	0.275	-0.019	0.464
4	3.912	3.920	0.285	0.008	0.204
5	3.401	3.441	0.277	0.040	1.176

3 结论

湖泊水质参数既具有随机性又具有结构性的特征,决定了可以应用地统计学的理论和方法来研究湖泊水质参数.在湖泊水质参数的研究中,地统计学中的克立格法是空间最优估计方法,另外还提供了衡量估计精度的指标——克立格方差.克立格方差的大小体现了估计精度的大小.

湖泊水质参数区域化变量空间最优估计

表 5 区域化变量 (Z_1) 靠近湖岸线和湖中岛
估计结果检验

Table 5 The tests in estimated Kriging accuracy of
regionalized variables (Z_1) near bank
and islands of Lake Tai

点号	实测值	估计值	克立格 方差	误差	相对误 差/ %
1	4.190	3.345	0.363	- 0.845	20.167
2	2.485	3.266	0.373	0.781	31.429
3	2.639	3.318	0.378	0.678	25.692
4	3.555	3.155	0.366	- 0.400	11.252
5	2.565	3.152	0.369	0.587	22.885
均值	3.087	3.247	0.370	0.160	22.285

表 6 区域化变量 (Z_1) 在湖中心估计结果检验

Table 6 The tests in estimated Kriging accuracy of
regionalized variables (Z_1) in the center of Lake Tai

点号	实测值	估计值	克立格 方差	误差	相对误 差/ %
1	3.178	2.886	0.359	- 0.292	9.188
2	3.258	3.010	0.359	- 0.248	7.612
3	3.091	3.278	0.359	0.187	6.050
4	3.258	3.100	0.363	- 0.148	4.543
5	2.996	3.357	0.359	0.361	12.049
均值	3.156	3.126	0.360	- 0.030	7.888

的估计精度是由体现区域化变量 $Z_{(x)}$ 空间结构特征的变差函数 $\gamma(h)$ 、信息采样点 x_i 与待估点

x_0 之间的空间配置关系、各信息采样点 x_i 相互之间的空间配置关系以及相应的变差函数值所决定。另外,也受湖岸线、湖中岛等因素的影响。空间结构直接决定着估计精度的大小,而与各信息采样点 x_i 值的大小无关,因而可在知道待估点实测值之前便可求出克立格方差。

参考文献:

1 王政权. 地统计学及在生态学中的应用[M]. 北京:科学出版社,1999. 102 ~ 149 .
2 A. G. 儒尔奈耳, CH. J. 尤日布雷格茨著. 侯景儒, 黄竞先, 杨尔煦等译. 矿业地质统计学[M]. 北京:冶金工业出版社,1982. 300 ~ 303 .
3 孙洪泉. 地质统计学及其应用[M]. 徐州:中国矿业大学出版社,1990. 96 ~ 125 .
4 中国地质学会数学地质专业委员会主编. 中国数学地质(2)[M]. 北京:地质出版社,1990. 12 ~ 16 .
5 张征,程祖锋,王恩祥等. 岩石参数随机场空间最优估计精度分析与特异值研究[J]. 岩土工程学报,1999, 21 (5):586 ~ 590 .
6 张征,刘淑春,邹正盛,等. 岩土参数的变异性及其评价方法[J]. 土木工程学报,1995. 28(6), 43 ~ 51 .
7 王仁铎,胡光道. 线性地质统计学[M]. 北京:地质出版社,1988. 68 ~ 71 .
8 M. 戴维著. 孙惠文,刘承祚译. 矿产储量的地质统计学评价[M]. 北京:地质出版社,1989. 289 ~ 294 .

欢迎订阅《环境污染治理技术与设备》

该刊由中国科学院生态环境研究中心主办,创刊于1980年,原名为《环境科学丛刊》,后更名为《环境科学进展》。经科技部和中国科学院批准,2000年1月起改名为《环境污染治理技术与设备》。该刊从2002年1月起将由双月刊改为月刊。

该刊主要报道国内外环境污染控制技术与设备的最新研究与开发的进展、成果和发展趋势,以及相关政策法规和国内外有关信息,并以较大篇幅刊登污染治理技术与设备方面的广告。

该刊的办刊宗旨是:为从事环境污染治理技术研究的科研人员和污染治理设备的生产开发商及用户,架起信息交流与传播的桥梁,积极推动我国环境污染治理技术与设备的研究开发和推广应用。

该刊的读者对象为:环境科研人员、环境管理人员、环保企业界人士、大专院校师生和社会其他各界人

士。

主办:中国科学院生态环境研究中心
编辑:《环境污染治理技术与设备》编辑部
地址:北京市海淀区双清路18号(北京2871信箱)
邮编:100085
电话:(010)62941074
传真:(010)62941074
E-mail:yzhdl074@263.net
国内发行:北京报刊发行局
国外发行:中国出版对外贸易总公司
国际标准刊号:ISSN 1008-9241
国内统一刊号:CN 11-4391/X
邮发代号:82-448
定价:12元/期