

基于 GIS 空间分析的海底表层沉积物粒度分布特征插值研究

严立文^{1,2,3}, 黄海军^{1,2}, 刘艳霞^{3,4}

(1. 中国科学院 海洋研究所, 山东 青岛 266071; 2. 中国科学院 海洋地质与环境重点实验室, 山东 青岛 266071; 3. 中国科学院 研究生院, 北京 100039; 4. 中国科学院 烟台海岸带研究所, 山东 烟台 264003)

摘要: 为了对比研究不同空间插值方法在沉积物粒度空间分布表述上的精度及适用性, 在 2004 年 11 月荣成宁津小型海湾黑泥湾综合调查的基础上, 探讨了表层沉积物粒度参数的空间插值, 比较了 IDW(反距离加权法)、Kriging(克里金插值法)、Spline(样条插值法)与 NN(自然邻域法)4 种 GIS 空间分析方法的特征、差异及实效性, 并对影响插值结果准确性的因素进行分析。结果显示, 在黑泥湾表层沉积物粒度结果的空间插值计算中, 从插值准确性和空间表达能力两方面考虑, IDW, Spline, Kriging 和 NN 中以 IDW 法较为适宜, 但要考虑到“牛眼效应”的出现会与局部实际情况存在差异; Kriging 法和 NN 法的插值结果准确性较高, 但其空间表达能力稍逊; Spline 法在近岸带表层沉积物粒度特征插值中的应用性相对较差。海洋调查要素空间分布的内在规律性是控制插值结果的主要因素, 表层沉积物分布以长期稳定的潮流、地形特征为主导因素, 呈现由岸向海条带状分布的特征。数据均匀分布区域的插值结果要优于边界区和突变区; 另外, 在选取的 3 个观察尺度上, 不同插值方法的误差均与野外取样网格间距呈显著线性正相关。

关键词: 黑泥湾; 表层沉积物; 地理信息系统; 插值方法

中图分类号: P736.2

文献标识码: A

文章编号: 1000-3096(2010)04-0058-07

海底沉积物的粒度特征是海洋沉积学研究的一项重要指标, 其中蕴含着海洋动力过程与机制^[1,2]、气候变化^[3,4]以及生态系统演化^[5,6]的大量信息。但是大范围高精度的海洋区域调查成本过高, 实施起来难度很大, 在空间布点取样调查的基础上进行空间拟合插值是当前的技术手段下获取沉积物粒度空间分布趋势的有效方法^[7]。因此, 如何利用有限的观测数据, 借助于空间插值方法估计其合理的空间分布具有重要的实际意义。

空间插值是场态离散数据处理常用的方法之一, 通过离散的已知数据或分区数据推求出任意点或分区数据, 可以生成或补充原始数据的不足, 使原始离散分布的数据转换成规则的格网数据, 具体可分为空间数据的内插与空间数据的外推^[8]。目前空间插值在海洋调查要素(水深、表层沉积物和悬浮体等)的等值线自动制图、数字高程模型(DEM)的建立、不同区域平面特征分析和界线相关分析和比较研究等方面已有广泛的应用。但对空间插值而言, 没有绝对最佳的方法^[9], 在复杂的区域条件下如何利用有限的实测数据得到相对准确的插值结果是空间插值研究中的一个难题, 里面涉及到原始数据特征、插值区域的独特性以及插值模型的精度和适应性。不同的插值方法和参数的选择可能会获得差异较大的结果。因此, 在对某一区域表层沉积物粒度

参数进行插值操作之前, 对不同插值方法进行比较和评估是很有必要的。

作者以 2004 年宁津附近海域综合调查时取得的黑泥湾表层沉积物为样本, 探讨了其粒度数据的空间插值情况。对 Arcgis 中的 3D Analyst 模块中的反距离加权法(IDW)、克里金法(Kriging)、样条法(Spline)和自然邻域法(NN)4 种插值方法的结果进行了对比和分析, 评价了不同方法插值的相对误差, 并对影响插值结果准确性的因素进行了分析, 可以为实现近岸海底表层沉积物空间分布特征的正确提取提供依据。

1 材料与方 法

1.1 研究区概况与调查取样

黑泥湾是山东半岛东端的小型近岸开敞型海湾

收稿日期: 2007-10-18; 修回日期: 2009-11-18

基金项目: 中国科学院海洋地质与环境重点实验室开放基金项目(MGE2009KG04); 我国近海海洋综合调查与评价专项项目(SD-908-01-03-01)国家自然科学基金项目(40676037)

作者简介: 严立文(1979), 男, 山东沂水人, 助理研究员, 从事海岸带环境遥感与 GIS 研究, 电话: 0532-82898935, E-mail: yanliwen@qdio.ac.cn

(图 1), 研究海区的水深 ≤ 30 m, 等深线基本与岸线平行。老炕石——黑石礁附近有水下暗礁分布, 等深线较为密集, 其南北两侧为浅湾, 湾内等深线稀疏, 地形平缓。由于山东半岛东端突入黄海, 海洋动力作用较强, 研究海区潮汐类型为不正规半日潮, 实测涨潮流为西南向, 最大流速 107 cm/s, 落潮流方向为东北, 最大流速为 105 cm/s^[10], 以上地形地貌特征与水动力条件控制着本区海底表层沉积物分布的空间格局。

定量评价。其中在老炕石——黑石礁一线的南、北部海湾各布设 3 个校验站位, 另外在黑石礁、楮岛与镆钁岛东侧的外围测线中各选取 1 个站位, 可以代表表层沉积物在研究区不同位置、水下地形与水动力条件下的分布情况。

1.2 数据分析方法

为了更加准确地重建黑泥湾海区表层沉积物的空间分布趋势, 以便进一步进行水动力条件特征和沉积环境演化的分析, 有必要对多种数据变换和空间插值方法进行比较, 找到更适合插值要素特征和区域特征的方法。在表层沉积物粒度参数计算的基础上, 通过 ESRI 开发的 Arcgis 软件 3D analyst 模块中的 IDW, Kriging, Spline 和 NN 4 种方法进行插值, 获得了不同插值方法得到的几种粒度特征结果, 并进行比较分析。

IDW 法是最常用的空间内插方法之一, 以插值点与样本点之间的距离为权重进行插值, 插值点越近的样本点赋予的权重越大, 其权重贡献与距离成反比。本文采用插值尺度为 46.5 m, 权重为 2。

Spline 法是用一种数学函数来估计值, 最小化所有的表面曲率, 逼近曲面的一种方法, 易操作且计算量不大。实际插值尺度为 46.5 m, 权重 0.1。

NN 法与 IDW 法类似, 也是一种权重插值方法, 但其是基于不规则三角网 (Delauney triangulation) 的最邻近结点插值, 适用于非均匀数据插值, 插值过程中无需进行插值范围、邻点数和权重等参数定义。本文中实际插值尺度为 105.87 m。

Kriging 法又称为空间自协方差最佳插值法, 由南非地质学家 Krige 于 1951 年提出, 1962 年法国学者 Matheron 引入区域化变量概念, 进一步推广和完善了克里金法。克里金法的优点是以空间统计学为坚实的理论基础, 物理含义明确; 不但能估计测定参数的空间变异分布, 而且还可以估算估计参数的方差分布; 通过权重系数的确定, 使内插函数处于最佳状态, 即对给定点上的变量值提供最好的线性无偏估计。其缺点是计算步骤较为繁琐, 计算量大, 且变异函数有时需要根据经验人为选定。文中选取普通克里金法采用球状模型 (Ordinary-Spherical) 进行插值, 插值尺度为 46.5 m。

在以上插值分析的基础上, 采用相对误差中值 (median relative error, E_{mr}) 和均方根误差 (root

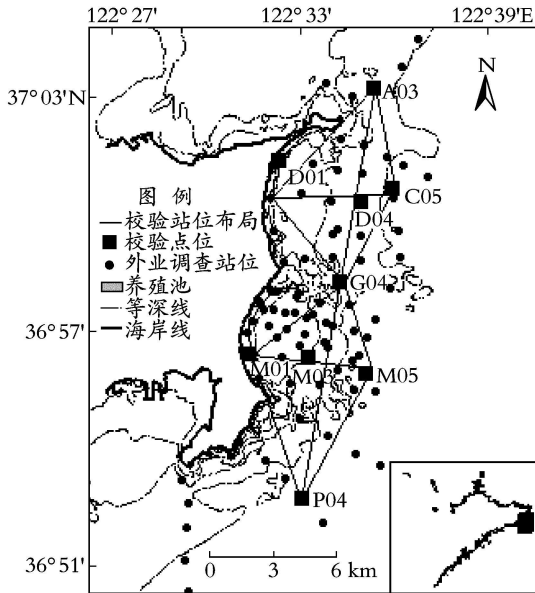


图 1 荣成海泥湾取样站位分布及校验点布置

Fig. 1 Sampling stations of surface sediments and validating position settings

2004 年 11 月份对宁津近岸黑泥湾海区的近岸表层沉积物、悬浮体、水流等海洋要素展开综合调查, 以获得近岸海区的沉积环境演变及水沙运移特征。根据调查要求, 取样站位以黑泥湾顶为中心呈辐射状分布, 湾内较为密集, 湾外站位布置相对稀疏, 在湾顶的凹进位置由于人工圈围养殖和暗礁的存在未进行取样。累计布置测线 17 条, 获得表层沉积物样品 75 个 (图 1)。样品的粒度测试在中国科学院海洋研究所完成, 测试仪器为法国产 Cilas 940L 型激光粒度仪, 测量粒径范围为 $0.3 \sim 2000$ μm , 多次重复测量的误差小于 3%。表层沉积物按照 Shepard 分类法^[11]进行分类, 并计算中值粒径、分选系数等粒度参数。

根据黑泥湾的地形特征和取样站位分布情况, 均匀选取 9 个站点作为检验数据, 把不同插值方法得到的结果与实测值进行比对分析, 并对误差进行

mean squared error, E_{rms}) 分析的方法对校验站位插值结果进行评价, E_{mr} 可以反映插值主体部分的相对准确程度, 而 E_{rms} 则可以反映插值的灵敏度和极值效应, 并评价不同插值方法在局部数据突变区插值的适应能力。

2 结果与讨论

2.1 黑泥湾海区表层沉积物分布特征

研究海区不同站位表层沉积物粒度参数的统计结果显示, 不同样品中值粒径的最大值为最小值的 7.3 倍, 各样品的相对标准偏差为 30.7%; 不同样品分选系数的最大值为最小值的 10.1 倍, 各样品的相对标准偏差为 40.8%。另外, 各个表层沉积物样品分选系数均小于 2.2, 分选程度为中等及以上, 平均为 1.0, 表明海区表层沉积物经受的水动力作用充分, 空间分布基本稳定平缓。海区沉积物的空间分布特征会对空间数据插值的精度产生一定影响。

黑泥湾海区校验站位粒度参数的计算结果见表 1。D01, M01 两个校验站点均位于近岸区, 而 M05 站位附近存在一小片砂区, 位于砂-黏土质粉砂的过渡区, 其余 6 个站位为均匀分布。不同站位的粒度参数有平行岸线分带的趋势, 中值粒径由岸向海增大, 表明表层沉积物的中值粒径逐渐减小; 分选系数由岸向海逐渐增大, 分选性由很好过渡为好。通过对以上代表性校验站位进行实测值与插值结果的比对分析, 可以探讨不同插值方法在本海区的适用情况, 并进一步分析影响插值结果精度的因素。

表 2 不同 arcgis 空间分析方法插值的相对误差结果

Tab. 2 Interpolation errors of different GIS methods

样品号	中值粒径				分选系数			
	IDW	spline	kriging	NN	IDW	spline	kriging	NN
A03	0.13Φ	0.29Φ	0.12Φ	0.29Φ	0.08	0.41	0.14	0.32
C05	0.02Φ	0.11Φ	0.07Φ	0	0.03	0.01	0.06	0.04
D01	0.74Φ	0.03Φ	0.74Φ	0.14Φ	0.78	0.24	0.49	0.18
D04	0.01Φ	0.13Φ	0	0.06Φ	0.08	0.08	0.07	0.03
G04	0.12Φ	0.03Φ	0.15Φ	0.06Φ	0.26	0.05	0.2	0.18
M01	0.62Φ	0.33Φ	0.79Φ	0.15Φ	0.05	0.18	0.15	0.34
M03	0.02Φ	0.08Φ	0.02Φ	0.03Φ	0.12	0.01	0.11	0.12
M05	0.18Φ	0.42Φ	0.17Φ	0.31Φ	0.08	0.66	0.19	0.05
P04	0.24Φ	0.14Φ	0.21Φ	0.22Φ	0.19	0.19	0.21	0.29
E_{mr}	0.13Φ	0.13Φ	0.15Φ	0.14Φ	0.08	0.18	0.15	0.18
E_{rms}	1.11Φ	1.29Φ	1.17Φ	1.08Φ	0.21	0.34	0.2	0.22

表 1 校验站位表层沉积物粒度参数的实测结果

Tab. 1 Grain size data in sampling positions for validation of interpolation

样品号	中值粒径	分选系数
A03	6.2Φ	1.2
C05	6.3Φ	1.4
D01	2.6Φ	0.5
D04	5.8Φ	1.3
G04	6.0Φ	1.4
M01	2.5Φ	0.7
M03	5.7Φ	1.3
M05	7.1Φ	1.3
P04	6.3Φ	1.4

2.2 插值方法的比较分析

表 2 即为 9 个校验站位插值结果与实测值之间的误差分布。中值粒径采用以上不同 4 种方法的总体差异性相对较小, E_{mr} 在 13% ~ 15% 的范围内, 表明在大部分区域里, 4 种插值方法得到的误差接近; 而 E_{rms} 值则存在差别, 其中 NN 法为最低, E_{rms} 为 1.08, IDW, kriging 和 spline 法次之, 分别为 1.11, 1.17 和 1.29。就中值粒径而言, 不同插值方法的误差分布存在差异。NN 法插值结果的误差分布最为均匀, IDW 与 kriging 法的插值误差大小与范围均较小。spline 法插值结果的 E_{mr} 较小, 但 E_{rms} (1.29) 为最高, 表明 spline 的插值误差分布范围较大, 其在某些站位的插值结果可能存在较大偏差。

不同方法对分选系数的插值误差 E_{mr} 的差异性较大, IDW 的 E_{mr} 最小为 8%, kriging 法次之为 15%, 与中值粒径不同的是, 分选系数的 NN 插值结果的 E_{mr} (18%) 为最大。这表明, 从大部分数据分选系数的插值结果而言, IDW 的插值精度为最高。 E_{ms} 的结果与 E_{mr} 相类似, IDW 与 kriging 方法的插值误差范围均为较小。

由以上分析可以看出, 以上 4 种插值方法中并没有哪种具有显著的优势, 即使同种插值方法在不同站位之间插值的精度也会不同。插值方法的适用性还要根据具体的插值要素和插值站位来判别, 并不存在一个普适的方法。如果不考虑插值方法对粒度参数空间分布趋势的表达能力的单就其总体精度而言, IDW 的插值精度最高, Kriging 和 NN 法次之, 而 Spline 法的插值精度相对较低。现有的部分研究结论^[12,13]认为 Kriging 方法以其变异函数的选取和调试更具有适应性, 准确性更高, 但在本文中的实例并不明显。这是因为 Kriging 插值方法对原始数据

有较高的要求, 一般在呈正态分布的条件下会得到较高的插值精度, 但野外的实测数据很难严格呈正态分布, 这就影响了 Kriging 法的适用范围。在对原始数据进行插值之前, 预先对原始数据的规律性进行分析是很有必要的。

不同插值方法获得的结果见图 2。4 种插值结果的条带状分布明显, 且近岸区条带较为密集, 向外过渡到 -20 m 等深线处较为稀疏, 这是对海底表层沉积物空间分布规律的真实反映。表层沉积物的空间分布主要受区域长时间尺度上潮汐、潮流以及波浪作用的影响, 在缺乏人为扰动的情况下, 其粒度、分选系数等分布特征可以表征该区域水动力条件的长期稳定形态, 短时间的局部干扰对其影响较小。因此表层沉积物插值的内在控制机理较为单一、明确, 主要可以体现海区长时间尺度上的稳定流态^[14]。总体而言, 海底表层沉积物粒度特征的空间插值结果的误差较小, 不同插值方法的总体误差中值 E_{mr} 为 14%。

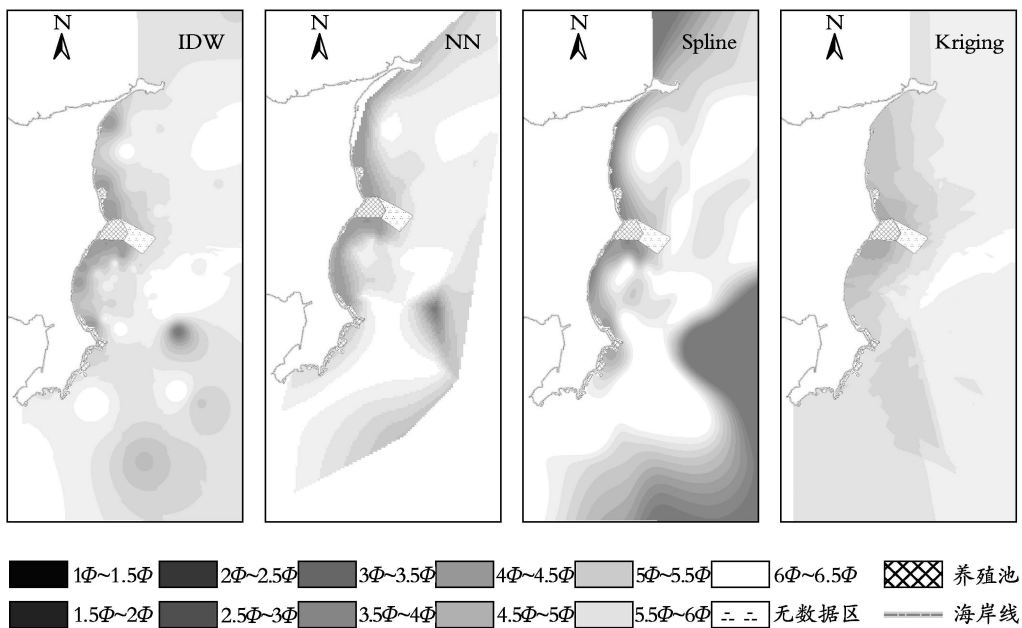


图 2 黑泥湾海区表层沉积物中值粒径的插值结果

Fig. 2 Spatial distributions of mean grain size data based on different interpolation methods

海洋要素空间数据分析的一个重要目的是探讨插值要素空间分布特征的规律性。以上插值方法在空间表达能力上具有较大的差异。IDW 的空间表达能力最好, 类型过渡比较真实, 其带状分布规律也得到很好的体现。其不足之处是极值分布区“牛眼”效应过于明显, 会影响到插值的局部准确度。Kriging 和 NN 法的插值结果在不同粒径等值线的过渡略显凸凹, 并且 NN 法的空间插值范围不能预先设定, 会

在边缘区存在“空白地带”。而 Spline 法的插值结果则过于柔和, 概化抽象的过多, 且根据大量插值实验的结果, 其形成的插值数据会溢出原始数据区间, 带来数据冗余和误差。

综上所述, 从局部数据的准确度和空间表达能力两方面考虑, 在黑泥湾海区粒度参数的 Arcgis 插值分析中以 IDW 法相对为最优, Kriging 和 NN 法次之, Spline 法在近岸粒度特征插值上的应用则值得探讨。

2.3 不同点位的表层沉积物中值粒径误差分布差异

以上4种插值方法在不同点位的插值误差具有明显差异。为探讨这种区位性差异,按照实测数据的位置和粒度特征将黑泥湾海区分别定义为均匀区、边界区和突变区。边界区是指临岸的取样区域,均匀区是指海湾内部及外侧表层沉积物类型均匀过渡的区域,黑泥湾海区大部分定义为均匀区;而突变区是指沉积物粒度特征强烈转变的区域。由图3可见,近岸区(D01, M01)的插值误差最大,不同插值方法的总体误差范围为37.2%±27.7%,突变区(M05)附近次之,为25.8%±20.3%,而均匀插值区得到的误差结果最小,误差范围仅为12.1%±9.7%。以上结果说明,插值点的区位和原始数据的均一程度对插值误差的分布也有重要影响。

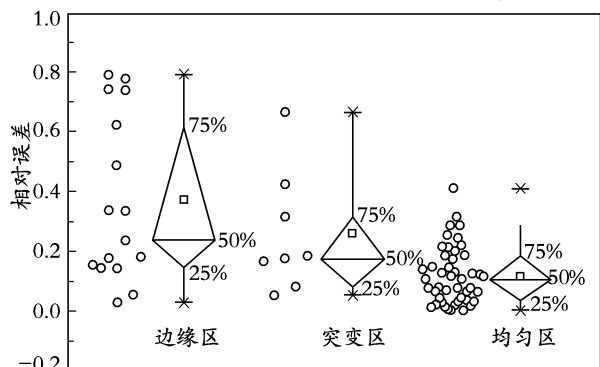


图3 边界区、均匀区和突变区的插值误差分布

Fig. 3 Interpolation errors in different deposit environments
GIS空间插值方法中的数学模型是建立在现有

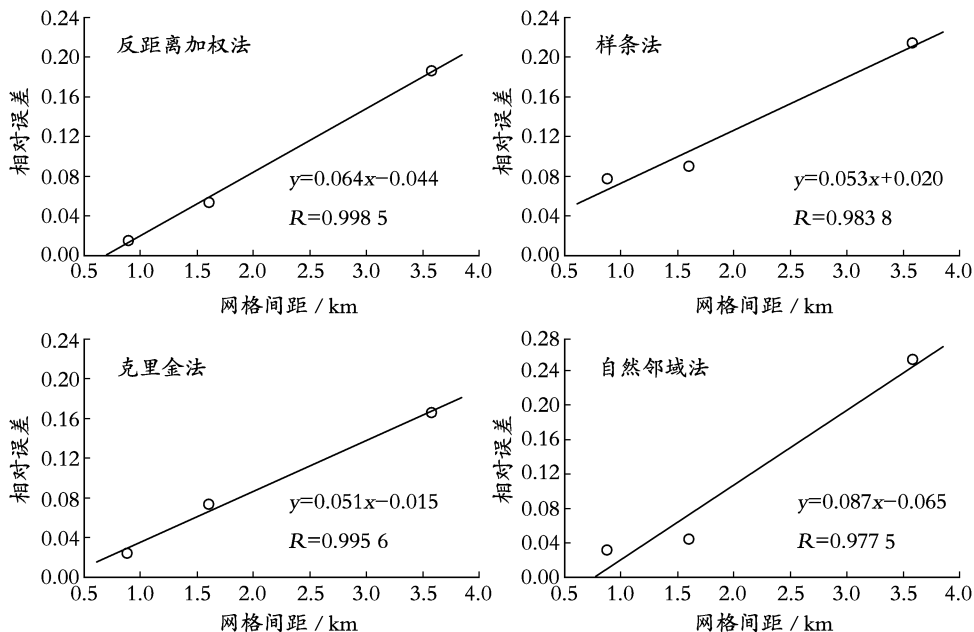


图4 取样网格间距与插值相对误差的关系

Fig. 4 Correlation between interpolation errors and gridding sizes

数据的基础上,对可能存在的极值难以预估,而极值周围的突变用常规插值方法获得的结果也不理想,这是各种插值方法普遍存在的一个问题,在实际分析中必须根据数据的空间特征建立适合的突变区插值方法。边界效应是不均一问题的一个特殊类型,即误差的高值区域集中在样区的边界,样区边界精度衰减较内部更为迅速。其原因在于边界区域参与空间插值的样点个数减少,且对待估点存在一定的方位影响,这种用于插值的原始数据空间分布的不均一性直接导致了用于插值的有效信息的缺失,使插值结果带有较大误差^[15]。而从表层沉积物的空间分布规律来看,近岸带为波浪强作用带,近岸区的沉积物粒度特征由岸向海变化较大,反应在插值结果上必然是较大的误差。

2.4 不同观测网格分辨率下的插值误差

根据校验站位与周围点位的间距不同,以1000 m和3000 m为界将网格间距划分为大、中、小3种,网格间距采用Le Roux 5点法^[16]定义,即为校验点与周围邻近4点的平均距离。检验站位实际取样网格的平均间距分别为881, 1604和3583 m。如图4显示,插值误差与取样网格间距呈显著线性正相关关系。尤以IDW法和Kriging法最为显著,相关系数分别为0.9985和0.9956,在95%的置信水平下显著正相关;Spline法和NN法在85%的置信水平下呈正相关。这说明观测网格的大小对插值结果会有直接的影响,本文提到的4种GIS插值方法都存在这一现象。

3 结论

基于 GIS 3D 空间分析模块对插值样本区——黑泥湾的表层沉积物粒度参数进行空间插值, 针对不同空间插值方法进行分析评价并就影响插值结果因素展开探讨, 可以得出以下结论: (1) 对近岸海区表层沉积物粒度参数的空间插值而言, 从插值准确性和空间表达能力两方面考虑, IDW, Spline, Kriging 和 NN 法中以 IDW 法较为适宜, 但要考虑到“牛眼效应”的出现会与局部实际情况有差异; Kriging 法和 NN 法的插值结果准确性较高, 但其空间表达能力稍逊; Spline 法的插值应用性相对较差。(2) 相对不同插值方法选取对插值结果造成的影响, 实际取样站位的空间分布特征对插值结果的影响更大。因为无论哪种插值方法的插值均建立在相邻网格的影响的基础上, 分析结果表明, 对于选取的 3 个观测尺度, 插值误差与实际取样网格的大小呈显著正相关。(3) 在近岸区表层沉积物粒度参数的空间插值研究中, 数据均匀分布区域的插值结果明显优于突变区和边界区。“边界效应”是一个不容忽视的问题, 一方面与近岸海区沉积物受波浪、潮流等较强动力条件作用, 沉积物类型变化较大有关; 另外, 插值数据的不均匀分布也是边界效应的重要原因。在空间插值中如何消除边界效应是个值得关注的问题。

参考文献:

[1] 高抒, 李安春. 浅海现代沉积作用研究展望[J]. 海洋科学, 2000, 24(2): 1-3.

[2] 朱玉荣. 末次盛冰期以来渤、黄、东海陆架底质分布演变过程模拟研究[J]. 黄渤海海洋, 2002, 20(1): 11-24.

[3] 李雪铭. 辽南滨海黄土的粒度分析和古气候[J]. 海洋科学, 1991, 2: 33-37.

[4] 肖尚斌, 李安春, 蒋富清. 近 2 ka 来东海内陆架的泥质沉积记录及其气候意义[J]. 科学通报, 2004, 49(21): 2 233-2 238.

[5] Boatman C D, Murray J W. Modeling exchangeable NH_4^+ adsorption in marine sediments: process and controls of adsorption [J]. **Limnology and Oceanography**, 1984, 27: 99-110.

[6] 陈木宏, 郑范, 陆均, 等. 南海西南陆坡区沉积物粒级指标的物源特征及古环境意义[J]. 科学通报, 2005, 50(7): 684-690.

[7] 贾建军, 程鹏, 高抒. 利用插值试验分析采样网格对粒度趋势分析的影响[J]. 海洋地质与第四纪地质, 2004, 24(3): 135-141.

[8] 邬伦, 刘瑜, 张晶, 等. 地理信息系统原理方法和应用[M]. 北京: 科学出版社, 2001. 178-192.

[9] 李新, 程国栋, 卢玲. 空间内插方法比较[J]. 地球科学进展, 2000, 15(3): 260-265.

[10] 山东省科学技术委员会. 山东省海岛志[M]. 济南: 山东科学技术出版社, 1995. 130-131.

[11] Shepard F P. Nomenclature based on sand-silt-clay ratios [J]. **Journal of Sedimentary Geology**, 1954, 24(3): 151-158.

[12] 靳国栋, 刘衍聪, 牛文杰. 距离加权反比插值法和克里金插值法的比较[J]. 长春工业大学学报, 2003, 24(3): 53-57.

[13] 刘登伟, 封志明, 杨艳昭. 海河流域降水空间插值方法的选取[J]. 地球信息科学, 2006, 8(4): 75-83.

[14] 陈吉余, 沈焕庭, 恽才兴. 长江口动力过程和地貌演变[M]. 上海: 上海科学技术出版社, 1988. 314-322.

[15] 于婧, 周勇, 聂艳. 江汉平原耕地土壤氮素空间尺度套合与变异规律研究[J]. 中国农业科学, 2007, 40(6): 1 297-1 302.

[16] Le Roux J P. An alternative approach to the identification of net sediment transport paths based on grain size trends [J]. **Sedimentary Geology**, 1994, 94: 97-107.

GIS-based analysis of spatial interpolations of sediments grain size in Heini Bay

YAN Li-wen^{1,2,3}, HUANG Hai-jun^{1,2}, LIU Yan-xia^{3,4}

(1. Institute of Oceanology, the Chinese Academy of Sciences, Qingdao 266071, China; 2. Key Laboratory of Marine Geology and Environment, the Chinese Academy of Sciences, Qingdao 266071, China; 3. Graduate University of the Chinese Academy of Sciences, Beijing 100039, China; 4. Yantai Institute of Coastal Zone Research, the Chinese Academy of Sciences, Yantai 264003, China)

Received: Oct., 18, 2007

Key words: Heini Bay; surface sediments; geographic information system (GIS); interpolation methods

Abstract: Interpolation applications on grain size parameters distribution of surface sediments in Heini Bay with GIS-based methods were discussed. Four interpolation methods (IDW, Kriging, Spline, NN) were compared and the factors influencing the interpolation accuracy were discussed from several aspects. The results indicate that considering interpolation accuracy and spatial expression IDW is relatively best among four interpolation methods in process of grain size parameter interpolation in Heini Bay. But “bull’s eye” effects should be considered in IDW. Kriging and NN methods have a high accuracy and relatively bad spatial description. The applications of Spline methods isn’t adaptive to grain size interpolation in nearshore sea. Compared with dissimilarity of different interpolation methods inner regularity, spatial distributions and sampling density contribute more in interpolation precision. The grain sizes have a thinning tendency seaward on control of steady currents and topography elements. The study also showed that interpolation errors are typically smaller in proportional spacing regions than that in border and abrupt-changing regions. The interpolation errors of different methods are all positive correlative to the sampling density in-situ.

(本文编辑: 刘珊珊)