

# 基于对数比转换方法的沉积物粒级组分空间预测与底质类型制图

刘付程<sup>1</sup>, 彭俊<sup>2</sup>

(1. 淮海工学院 测绘工程学院, 江苏 连云港 222005; 2. 盐城师范学院 城市与资源环境学院, 江苏 盐城 224002)

**摘要:** 以废黄河三角洲表层粒度分析数据为基础, 探讨了对数比转换和 kriging 插值相结合方法在沉积物粒级组分空间预测和底质类型制图中的应用。结果表明, 基于沉积物粒级组分原始数据的 kriging 预测方法难以保证各组分预测结果的非负和定和要求, 因而预测结果的可信度低; 而对数比转换 kriging 方法不但满足非负和定和要求, 而且还有着更优的组分预测结果和较高的底质类型制图精度。新方法对于开展定量化的沉积物粒级组分预测和底质类型制图具有参考价值。

**关键词:** 对数比转换; 沉积物粒级组分; 空间预测; 底质制图

中图分类号: P736.21 文献标识码: A 文章编号: 1000-3096(2015)06-0122-07  
doi: 10.11759/hyhx20130529001

底质类型图是海洋综合调查成果的基础图件之一, 在海洋工程、渔业、养殖、航海、国防以及海洋沉积环境研究等方面都有着广泛的应用价值<sup>[1]</sup>。利用沉积物粒度分析数据开展底质类型空间预测制图是一种常见的制图方式, 其技术途径可分为两类, 一是通过定量估计不同粒级组分(如砂、粉砂、黏土)的空间分布来进行底质类型识别和制图表达<sup>[2]</sup>, 二是通过采样站位的底质类型识别结果, 运用 Voronoi 图、指示 kriging 等方法来进行底质类型边界划分<sup>[3-4]</sup>。由于前者更符合底质制图的一般逻辑推理过程, 因而广泛被人们理解和接受, 但其关键是沉积物粒级组分的空间预测结果要可靠, 否则底质类型识别的准确性和制图成果的可信度将难以保证。然而在沉积物粒级组分的空间预测过程中, 人们往往忽视其作为成分数据所具有的特殊性, 这有可能导致与实际情况不相符的预测结果出现, 如同一位置处砂、粉砂和黏土组分的预测结果之和不能满足“定和”条件<sup>[5]</sup>。

成分数据是指具有“非负”和“定和”特性的一组数据, 也即各组分取值  $\geq 0$  且其加和结果为一常数<sup>[6-7]</sup>。成分数据的“定和”特性使得其各组分之间存在伪相关, 并由此产生“闭合效应”且不满足经典统计学方法的基本假设<sup>[6-9]</sup>, 因此盲目采用经典统计方法来处理成分数据有可能得出错误的结论<sup>[7]</sup>。沉积物粒级组分数据属于典型的成分数据, 在其空间预测过程中, “非负”和“定和”要求是评判预测效果的重要依据之

一, 也是后续开展底质类型识别和制图表达的前提条件, 然而在现有的多数文献报道中, 这一点往往被有意或无意地忽略了。

成分数据的“闭合效应”使得直接运用成分数据的原始值来开展组分的空间预测受到质疑<sup>[10]</sup>。著名的统计学家 Aitchison 教授提出了运用对数比转换方法来消除成分数据的“闭合效应”<sup>[8]</sup>。该方法将成分数据变换成其组分比值的对数, 从而实现了成分数据从单形空间向实数空间的映射, 并使得转换结果近似地服从正态分布<sup>[8-9]</sup>, 这为运用经典统计学方法来处理和分析成分数据创造了有利条件。本文以废黄河三角洲表层沉积物粒度分析数据为基础, 尝试运用对数比转换和 kriging 插值方法来对沉积物不同粒度组分的空间分布进行预测, 并评价预测结果的准确性, 在此基础上开展底质类型识别及制图表达。

## 1 材料与方 法

### 1.1 区域概况与数据

废黄河三角洲由 1128~1855 年间黄河南泛夺淮

收稿日期: 2013-05-29; 修回日期: 2013-07-27

基金项目: 国家自然科学基金项目(41306077); 淮海工学院自然科学基金项目(Z2014017)

作者简介: 刘付程(1971-), 男, 安徽安庆人, 副教授, 博士, 主要从事海岸环境演变和 GIS 与遥感应用方面研究, 电话: 0518-85895591, E-mail: iliufucheng@126.com

入海所带来的泥沙淤积而成,属于典型的淤泥质海岸。1855年黄河尾间北归后,因丧失了主要泥沙来源,三角洲被废弃转而进入了侵蚀调整阶段。废黄河三角洲近岸海域平均流速为 0.25 m/s,主流向为 NNW-SSE。据 1995~2005 年滨海海洋站波浪观测资料统计,研究区常浪向为 ENE,强浪向为 NE<sup>[11]</sup>。

2008 年 7~8 月,在废黄河口两侧布设了 13 条沉积物采样断面,断面从潮间带开始,向外海延伸 5~6 km。在 13 条断面上共采集表层沉积物样品 188 个(图 1)。沉积物样品经处理后采用 CoulterLS-100Q 型激光粒度分析仪进行粒度分析,获得全部样品的砂( $\Phi = -1\sim 4$ )、粉砂( $\Phi = 4\sim 8$ )和黏土( $\Phi > 8$ )组分的含量。

在全部 188 个样品中,随机抽取 50 个作为验证数据集,以进行沉积物粒度组分空间预测的准确性评价;其余的 133 个样品作为插值数据集用于沉积物粒级组分的空间预测(插值)。插值数据集和验证数据集的空间分布见图 1。

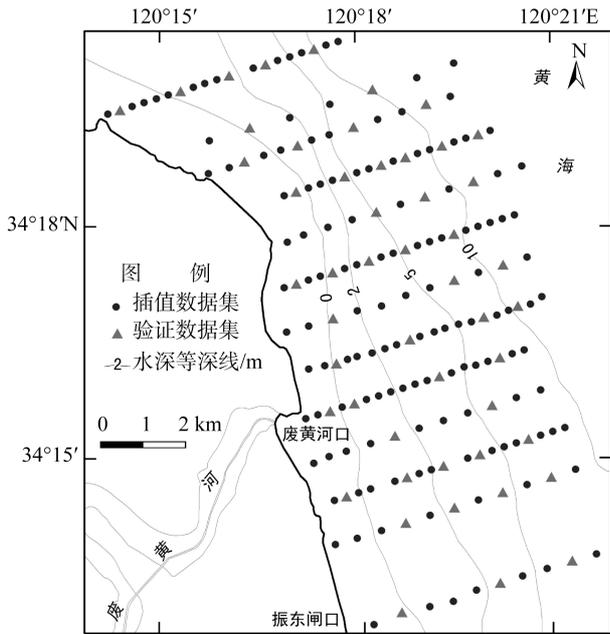


图 1 沉积物采样站位分布图

Fig. 1 The spatial distribution of sampling sites

### 1.2 对数比转换方法

成分数据对数比转换方法主要有加和对数比转换(additive log-ratio)、中心化对数比转换(centered log-ratio)和等角对数比转换(isometric log-ratio)等多种方法<sup>[8, 12]</sup>。近些年来,有学者为解决成分数据中的 0 值问题(因 0 值无法取对数)而对加和对数比转换方法进行了改进,提出了改进的加和对数比转换方法

(modified additive log-ratio)<sup>[10]</sup>。考虑到本次沉积物粒度分析数据中,部分样品的砂组分出现 0 值情况,因此本文采用改进的加和对数比转换方法对沉积物的砂、粉砂和黏土组分进行转换,其转换公式为:

$$y_{ij} = \ln \frac{x_{ij} + \eta_j}{\left( \prod_{j=1}^c (x_{ij} + \eta_j) \right)^{1/c}} \quad (1)$$

回转公式为:

$$x_{ij} = \frac{\exp y_{ij} - \frac{\eta_j}{1 + \sum_{j=1}^c \eta_j}}{\sum_{j=1}^c \exp y_{ij}} \left( 1 + \sum_{j=1}^c \eta_j \right) \quad (2)$$

式中,  $x_{ij}$  为第  $i$  个沉积物样品的第  $j$  种组分的含量;  $y_{ij}$  为第  $i$  个样品的第  $j$  种组分含量的转换值;  $\eta_j$  为常数,取样本中各组分数据除 0 以外最小值的一半;  $c$  为沉积物组分的类别数。

### 1.3 空间预测效果评价

由于沉积物的粒级组分数据为成分数据,因此对其空间预测效果的评价既要考虑各粒级组分空间预测结果的准确性,同时还要考虑各组分预测结果是否满足“非负”和“定和”的要求。

本文对不同粒级组分空间预测结果准确性的评价采用平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)和一致性指标(index of agreement,  $d$ )来进行。平均绝对误差和均方根误差越小、一致性指标越大表示预测效果越好<sup>[13]</sup>。各指标的计算公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Z_i - X_i| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_i - X_i)^2} \quad (4)$$

$$d = 1 - \left[ \frac{\sum_{i=1}^n (Z_i - X_i)^2}{\sum_{i=1}^n (|Z_i| + |X_i|)^2} \right], \quad 0 \leq d \leq 1 \quad (5)$$

式中,  $n$  为验证数据集中的样本数;  $X_i$ 、 $Z_i$  分别表示验证数据集中的第  $i$  个样本中某粒级组分的实测值和预测值;  $Z'_i = Z_i - \bar{X}$ ,  $X'_i = X_i - \bar{X}$ ,  $\bar{X}$  表示相应粒级组分实测值的平均值。

对于“非负”检验可直接由各粒级组分的预测结果来判别,而对“定和”的检验则是通过叠加各组分预测结果的栅格图,再逐个栅格判断其 3 组分预测值的加和是否为 100%。

## 2 结果与分析

### 2.1 沉积物不同粒级组分含量的描述性统计

从表 1 可以看出, 研究海域沉积物组分中粉砂的平均含量最高, 为 55.86%, 超出了黏土和砂的平均含量之和, 反映了废黄河三角洲沉积物的粉砂质特性。3 种组分的变异系数差别较大, 表明其空间分布的均匀性存在着显著的差异。砂的变异系数达到了 1.216, 说明其空间分布的异质性较强; 粉砂的变异系数最小, 反映其空间分布均匀性要比砂和黏土更为明显。表 1 中插值数据集与全体数据集有较为相似的统计参数, 可将其视为全体数据集的一个理想子集。由插值数据集开展沉积物粒级组分的空间分布预测结果能在一定程度上反映全体数据集所体

现的底质空间分布特征。

### 2.2 沉积物粒级组分含量的空间分布预测

沉积物粒级组分的空间分布预测通常是通过空间插值来实现的。由于沉积物不同粒级组分的空间分布与位置有关, 并表现出一定的空间自相关性, 因此地统计学中的 kriging 插值方法常被认为是一种理想的预测方法<sup>[14]</sup>。Kriging 插值方法本质上是一种加权平均方法, 其权重是在满足最优无偏估计条件下通过半方差函数求得的<sup>[15]</sup>。本文采用地统计学软件 GS+7.0 对经加和对数比转换后的砂、粉砂和黏土数据进行半方差计算和理论模型拟合, 模型参数由交叉验证法来确定。图 2 给出了相应的拟合理论模型及参数值(图 2a, b, c)。

表 1 沉积物粒级组分含量的统计特征

Tab. 1 Statistical characteristics of different grain size compositions of the sampling sediments

数据集	粒级组分	样本数	体积分数(%)				变异系数	偏态值	峰态值
			平均值	最小值	最大值	标准差			
全体数据集	砂	188	20.00	0.00	95.13	24.32	1.216	1.35	3.97
	粉砂	188	55.86	3.05	78.63	15.71	0.281	-1.57	5.16
	黏土	188	24.14	1.82	51.09	12.96	0.537	0.20	1.83
插值数据集	砂	138	21.19	0.00	95.13	25.08	1.183	1.29	3.87
	粉砂	138	55.22	3.05	78.63	16.29	0.295	-1.63	5.51
	黏土	138	25.59	1.82	51.09	13.26	0.518	0.14	1.77

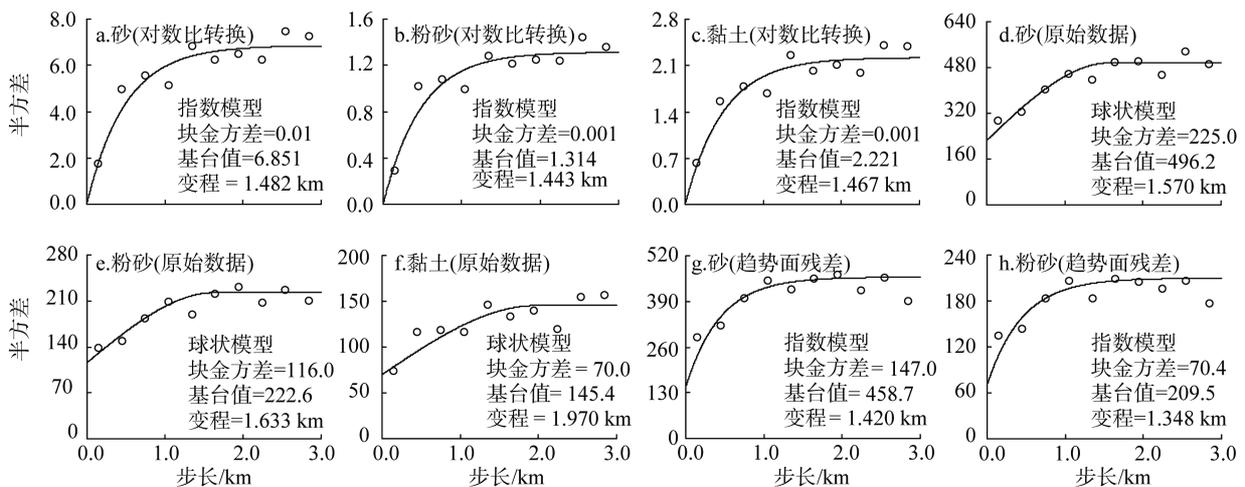


图 2 基于不同数据处理方法的沉积物粒级组分理论半方差模型

Fig. 2 The fitted models for smivariograms of sediment grain size compositions based on data processed with different methods

图 2 同时也给出了基于砂、粉砂和黏土原始数据的半方差散点图及其拟合理论模型参数(图 2d, e, f)。考虑到砂和粉砂原始数据的统计分布呈偏态(表 1), 其半方差函数可能不够稳健<sup>[15]</sup>, 有可能会影响

到砂和粉砂的最终插值结果, 因此, 对砂和粉砂原始数据进行了去除趋势化处理, 即将插值数据集中各样点砂和粉砂的实测值减去其各自的二阶趋势面(选择二阶趋势面是因为经检验发现砂和粉砂含量

的空间分布存在二阶趋势), 从而获得砂和粉砂的二阶趋势面残差。由于砂和粉砂二阶趋势面残差的统计分布均近似为正态, 由此计算得到的半方差函数更为稳健。图 2g, h 为砂和粉砂二阶趋势面残差的半方差函数图。

从图 2 可以看出, 经不同数据处理方法得到的砂、粉砂和黏土的变量半方差均可用指数模型或球状模型来拟合, 表明变量的空间变异具有明显的结构性特征<sup>[15]</sup>, 可进一步运用 ArcGIS 的普通 kriging 方法来对其空间分布进行插值, 插值栅格边长为 30 m×30 m。为便于对不同数据处理方法下的预测结果进行比较, 本文运用 ArcGIS 的栅格运算功能, 将所有插值结果回转到砂、粉砂和黏土的原始数据尺度, 即根据公式(2)将基于对数比转换数据的插值结果恢复到其相应的体积百分比数据, 得到基于对数比转换 kriging 方法的砂、粉砂和黏土空间预测结果分布图(图 3a, b, c); 将砂、粉砂二阶趋势面残差的 kriging 插值结果与其对应的二阶趋势面进行加和, 得到基于趋势面残差 kriging 方法的砂、粉砂空间预测结果分布图(图 3g, h)。图 3d, e, f 是基于砂、粉砂和黏土原始数据的 kriging 插值结果分布图。

从图 3 可以看出, 基于不同数据处理方法所获得的相应粒级组分的空间分布格局基本一致, 砂的

高值区主要分布在废黄河口南侧近岸区域, 呈与潮流方向一致的条带状分布; 粉砂的高值区主要分布在研究区域的西北部及东部; 黏土的低值区与砂的高值区基本对应, 其高值区主要分布在研究区域的北部近岸和东北部。基于对数比转换 kriging 方法的预测结果对 3 组分空间分布的刻画更为精细, 比较图 3a, b, c 和表 1 可以发现, 其预测结果的上、下限值与插值数据集基本接近。基于原始数据的 kriging 预测结果表现出显著的平滑效应, 其值域范围较插值数据集显著缩小(图 3d, e, f 和表 1)。趋势面残差 kriging 方法对砂组分的预测结果出现负值(主要出现在东北部采样点外侧区域), 而对粉砂的预测结果超出了实测数据值域范围的上限值, 与实测数据不符(图 3g, h 和表 1)。

### 2.3 沉积物粒级组分含量的预测结果评价

表 2 是根据验证数据集中各样点的预测值与实测值, 运用式(3)~式(5)计算得到的 MAE、RMSE、*d* 等评价指标的值。从表 2 可以看出, 砂、粉砂和黏土组分在不同预测方法下的 MAE、RMSE、*d* 值虽然差异不大, 但 3 组分的 MAE、RMSE 均表现为对数比转换 kriging 法<趋势面残差 kriging 法<原始数据 kriging 法, 而一致性评价指标 *d* 的排序正好相反, 这

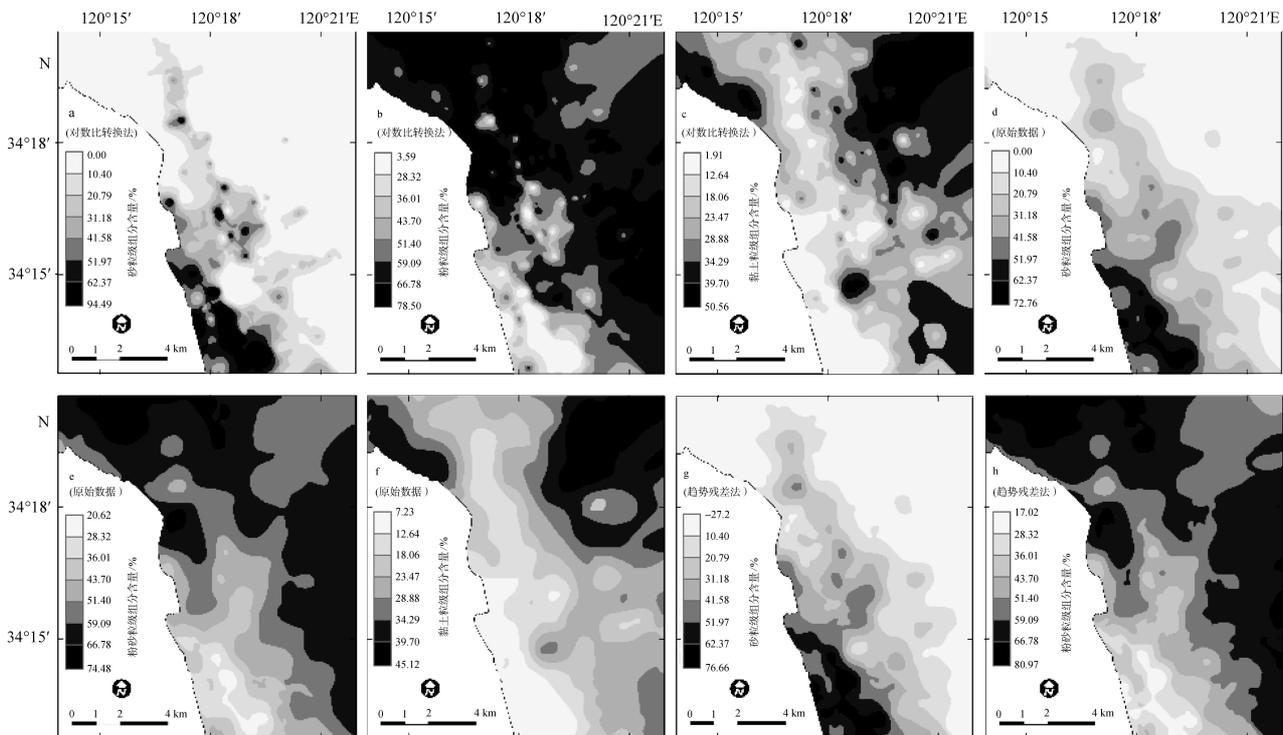


图 3 基于不同数据处理方法的沉积物粒级组分预测图

Fig. 3 Prediction maps of sediment grain size compositions based on different data processing methods

表 2 不同数据处理方法下的预测结果准确性评价

Tab. 2 Evaluation of prediction accuracy at validation sampling stations based on different data processing methods

数据处理方法	MAE(%)			RMSE(%)			<i>d</i>		
	砂	粉砂	黏土	砂	粉砂	黏土	砂	粉砂	黏土
对数比转换	9.30	7.50	6.90	13.52	10.02	8.39	0.858	0.837	0.813
原始数据	9.84	8.26	7.40	17.56	12.64	9.43	0.756	0.672	0.785
去除趋势面	9.38	7.87	—	13.68	10.56	—	0.828	0.746	—

表明仅从各粒级组分空间预测的准确性来说,对数比转换 kriging 法要优于另两种方法。

将同一方法的砂、粉砂和黏土的预测结果栅格图进行叠加,获得 3 组分“加和”分布图,可用来判别不同空间位置处 3 组分预测结果是否满足“定和”要求。图 4 给出了不同预测方法的 3 组分“加和”分布图,其中基于趋势面 kriging 方法的“加和”分布图中的黏土组分采用的是基于原始数据 kriging 方法的空间插值结果。从图 4 可以看出,基于对数比转换 kriging 方法的预测结果在所有空间位置处均满足“定和”为 100%的要求;而另两种方法只有不到 10%的区域面积满足这一条件,其绝大部分位置处的 3 组分预测结果之和均大于或小于 100%,表明其预测结果只具有空间格局表达的参考价值,不能作为定量分析的依据。

### 2.4 基于沉积物粒级组分预测结果的底质类型制图

考虑到只有对数比转换 kriging 方法对砂、粉砂和黏土的预测结果满足“非负”和“定和”要求,因此本文根据其预测结果,按照 Shepard 沉积物分类和命名方案,运用 ArcGIS 的栅格分析功能,逐个栅格判别其所属沉积物类型并最终绘制出研究区域的底质

类型分布图(图5)。从图中可以看出,研究海域存在 7 种底质类型,与实测数据的分类结果一致,其中黏土质粉砂的分布范围最广,占插值区域面积的 75%左右,主要分布在废黄河口北侧近岸和 5 m 等深线以深海域;其次是砂质粉砂,占插值区域面积的 11%,主要分布在废黄河口门外的邻近海域;粉砂质砂占总面积的 9%左右,主要分布在废黄河口南侧近岸区域;其他类型沉积物呈斑块状分布且面积均不大。

为进一步评价制图的精度,将制图结果与相应实测样本的沉积物类型进行比较,发现在全部 188 个采样站位处,共有 174 个站位的底质类型与制图结果是一致的,即底质制图的总体精度达到了 92.6%,其中插值数据集和验证数据集中的一致性样本数分别为 135 个和 39 个,占其各自样本总数的 97.8%和 78.0%,表明底质类型预测制图的总体效果较好。

### 3 结论

(1) 沉积物粒级组分数据属于成分数据,具有非负和定和特性。基于沉积物粒级组分原始数据的 kriging 预测方法难以保证预测结果的非负和定和要求,其预测结果只能在一定程度上反映组分的空间分布格局,不能作为定量分析的依据。而基于对数比

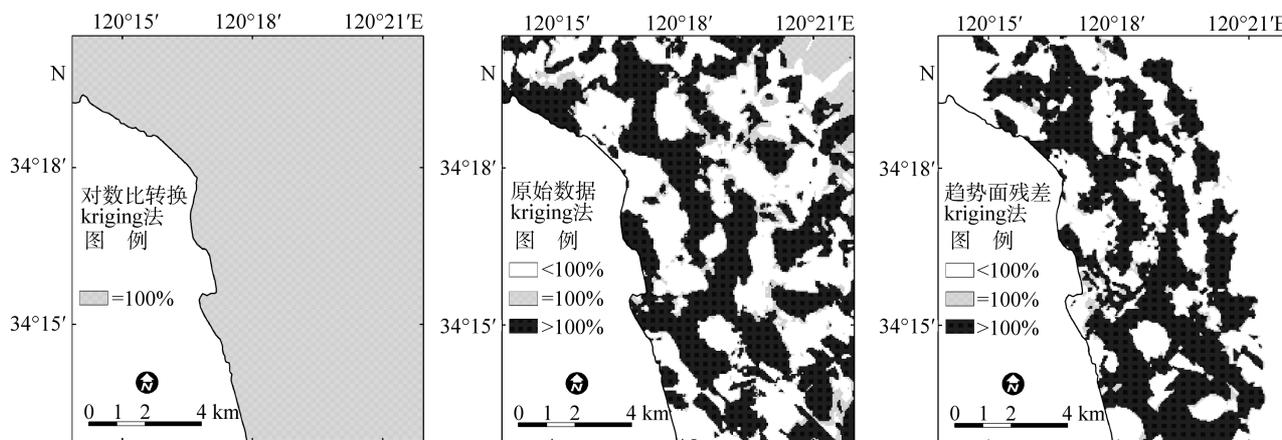


图 4 沉积物粒度组分预测结果加和分布图

Fig. 4 Spatial distribution of the sum of predicted results of sand, slit and clay

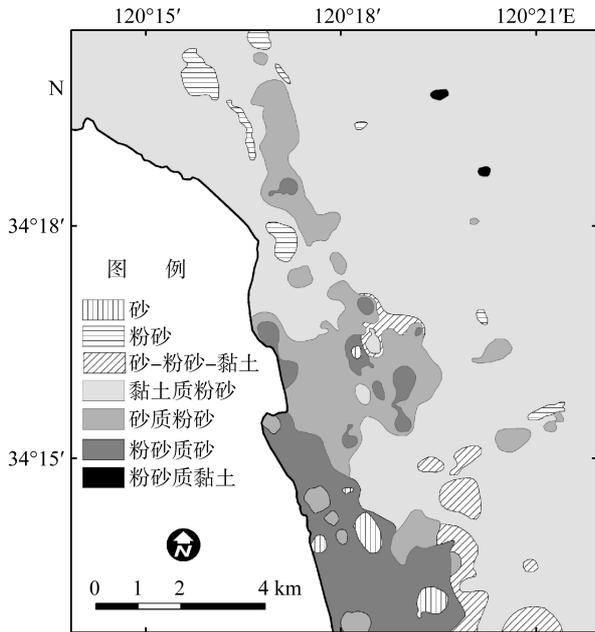


图 5 废黄河口海域底质类型分布图

Fig. 5 Sediment type map of the coast of the abandoned Yellow River Delta

转换 kriging 方法,不但能确保预测结果符合成分数据的基本要求,而且还有着较高的组分预测准确度,其预测结果可进一步用于底质类型识别和制图表达等定量分析过程。

(2) 基于对数比转换 kriging 方法的砂、粉砂和黏土预测结果,开展了废黄河三角洲海域的底质类型识别和制图表达。制图结果表明,研究海域存在 7 种底质类型,与实测数据的分类结果一致,其中以黏土质粉砂分布最广,占插值区域的 75%左右,且主要分布在废黄河口北侧近岸和 5 m 等深线以深海域;砂质粉砂主要分布在废黄河口门外的邻近海域,粉砂质砂主要分布在废黄河口南侧近岸区域,其他类型底质分布面积较小且呈斑块状零星分布。基于实测数据的评价结果表明,底质制图的总体精度达到了 92.6%,符合制图要求。

参考文献:

[1] 刘锡清. 最新中国近海陆架底质类型图[J]. 海洋地质与第四纪地质, 1992, 12(4): 11-20.

[2] 杨康, 张永战. 基于栅格叠合的沉积物底质图生成方法[J]. 第四纪地质, 2007, 27(5): 889-895.

[3] 王涛, 陈惠荣, 王少帅, 等. 基于 Voronoi 的海洋底质区域划界方法研究[J]. 测绘与空间地理信息, 2011, 34(2): 242-243.

[4] 刘付程, 彭俊, 张瑞, 等. 一种近海底质类型图生成的非参数方法[J]. 海洋通报, 2011, 30(5): 551-556.

[5] Jerosch K. Geostatistical mapping and spatial variability of surficial sediment types on the Beaufort Shelf based on grain size data[J]. Journal of Marine System, 2012, doi: 10.1016/j.jmarsys.2012.02.013.

[6] Aitchison J. The statistical analysis of compositional data[J]. J Royal Stat Soc B, 1982, 44(2): 139-177.

[7] 周蒂. 对数比统计分析及粒度数据中沉积水动力环境信息的萃取[J]. 沉积学报, 1996, 14(增刊): 149-157.

[8] Aitchison J. The statistical analysis of compositional data[M]. London: Chapman and Hall, 1986: 58-61.

[9] 周蒂. 地质成分数据统计分析—困难与探索[J]. 地球科学—中国地质大学学报, 1998, 23(2): 147-152.

[10] Odeh I O A, Todd A J, Triantafyllis J. Spatial prediction of soil particle-size fractions as compositional data[J]. Soil Science, 2003, 168 (7): 501-514.

[11] 陆勤. 废黄河三角洲淤泥质海岸稳定性研究[D]. 上海: 华东师范大学, 2011.

[12] Egozcue J J, Pawlowsky-Glahn V, Mateu-Figueras G, et al. Isometric logratio transformations for compositional data analysis[J]. Mathematical Geology, 2003, 35(3): 279-300.

[13] Willomtt C J. Some comments on the evaluation of model performance[J]. Bulletin American Meteorological Society, 1982, 63(11): 1309-1313.

[14] 刘付程, 张存勇, 彭俊. 海州湾表层沉积物粒度的空间变异特征[J]. 海洋科学, 2010, 34(7): 54-59.

[15] 王政权. 地统计学及其在生态学中的应用[M]. 北京: 科学出版社, 1999: 59-149.

# A logratio transformation based kriging method for spatial prediction of sediment grain size compositions and sediment type mapping

LIU Fu-cheng<sup>1</sup>, PENG Jun<sup>2</sup>

(1. School of Geodesy and Geomatics Engineering, Huaihai Institute of Technology, Lianyungang 222005, China; 2. School of Urban and Resources Environment, Yancheng Teachers University, Yancheng 224002, China)

**Received:** May, 29, 2013

**Key words:** logratio transformation; sediment grain size composition; spatial prediction; sediment type mapping

**Abstract:** The sediment grain size data are compositional data and are characterized by non-negative and constant sum. In spatial prediction of grain size compositions, one important aspect of the prediction quality is whether the prediction results meet the requirements of non-negative and constant sum, which is also the essential condition for sediment type recognition and mapping. In this paper, the grain size data obtained from the abandoned Yellow River Estuary are taken as an example to discuss the application of combined logratio transform and ordinary kriging in spatial prediction of grain size compositions and sediment type mapping. Results show that the kriging interpolation results by directly using the grain size data are unreliable owing to their dissatisfaction of non-negative and constant sum, while the prediction results obtained using the logratio-transformed grain size data and kriging interpolation method not only meet those two requirements, but also have a better prediction accuracy of grain size compositions and a relative high sediment type mapping precision. So the new method of combined logratio transform and ordinary kriging has reference value for quantitative sediment mapping.

(本文编辑: 刘珊珊)