# A non-parametric indicator Kriging method for generating coastal sediment type map

LIU Fu-cheng[1], PENG Jun[2], ZHANG Cun-yong[1]

1*. School of Geodesy & Geomatics Engineering, Huaihai Institute of Technology, Lianyungang* 222005*, Jiangsu Province, China;*

2*. State Key Lab of Estuarine and Coastal Research, East China Normal University, Shanghai* 200062*, China*

**Abstract:** Coastal sediment type map has been widely used in marine economic and engineering activities, but the traditional mapping methods had some limitations due to their intrinsic assumption or subjectivity. In this paper, a non-parametric indicator Kriging method has been proposed for generating coastal sediment map. The method can effectively avoid mapping subjectivity, has no special requirements for the sample data to meet second-order stationary or normal distribution, and can also provide useful information on the quantitative evaluation of mapping uncertainty. The application of the method in the southern sea area of Lianyungang showed that much more convincing mapping results could be obtained compared with the traditional methods such as IDW, Kriging and Voronoi diagram under the same condition, so the proposed method was applicable with great utilization value.

**Key words:** sediment type, non-parametric indicator Kriging, uncertainty, mapping

## 1 Introduction

Marine sediment type is the result of sediment classification and nomenclature according to the composition of different size fractions based on a certain kind of sediment classification system [1], and marine sediment type map is usually regarded as one of important fruits of marine sedimentary survey. Nowadays, marine sediment type map has widely practical values in modern sedimentary environment analysis, marine engineering and selection for anchoring sites of buoys and ships [2, 3]. Therefore, studying the mapping method of marine sediment types is of great importance in theory and practice.

The traditional marine sediment type map has usually been created based on the sediment types at the given sites, underwater topography information and expert knowledge, and such method can make good use of the known information and expert

Received on September 22, 2011

Corresponding author: iliufucheng@126.com

knowledge, but there also exists a certain degree of subjectivity and inefficiencies in the mapping process[2,4]. In recent years, with the rapid development of computer technology and spatial analysis theories, some data-driven mapping methods were also proposed in the field of sediment mapping. Zhang et al. (2007) presented a Voronoi diagram technique for generating sediment type map and discussed in detail the process of generating Voronoi diagram of sediment types [3]. Yang and Zhang (2007) introduced a raster overlay method to generate types and distribution of sediment map [4]. They firstly utilized spatial interpolation techniques such as Inverse distance weighting (IDW) and Kriging to obtain the spatial distribution maps of sand, slit and clay, and then overlaid these maps to get sediment type map through raster combination operations. These data-driven methods can avoid subjectivity in the mapping process to some extent, but there still exist some limitations in practices, for example, the method based on Voronoi diagram or IDW does not pays enough attention to spatial autocorrelation of sediments, and Kriging method requires the variables to meet second-order stationary or certain statistical hypothesis. These requirements have negative effects on the applicability of methods and the credibility of mapping results. Besides, most of the present methods still can not provide effective means to quantitatively evaluate the uncertainty of mapping results, which also restricts, to a certain extent, the effective application of mapping results. In this paper, a non-parametric indicator Kriging method was introduced for generating thematic map of costal sediment types, and the south sea area of Lianyungang was also taken as an example to test performance of the proposed method and evaluate the accuracy of its mapping results

## 2 Basic principle of non-parametric mapping method

### 2.1 basic idea of the method

Suppose there are M sediment samples got from a delimited sea area, and they are classified into N sediment types according to a certain kind of sediment classification system, such as Shepard's classification. At a given site, if the occurring possibility of the N sediment types are P1, P2, …, PN, respectively, there is a rational assumption: the sediment type at the given site could be considered as the type with the maximum of Pi (i = 1, 2, …, N). Based on the assumption, the key to the mapping process turns to be how to predict the occurring probability of each sediment type at any site in the delimited area, and then the sediment type can be determined by comparing the probability values.

If sampling process is regarded as a realization of spatial distribution probability of a specified type of sediment, the probability value at the sampling sites is either 1 or 0, where

1 represents the specified sediment type at a certain sampling site, otherwise it does not. Thereafter, a non-parameter geostaitsics method, namely indicator Kriging, can be used to predict the occurring probability of the specified type of sediments at all of the non-sampling sites.

## 2.2 Indicator Kriging

Indicator Kriging, firstly introduced by Journel[5], is a non-parametric form of conditional ordinary Kriging and usually applied to predict conditional cumulative distribution function of regionalized variable at any non-sampling sites. The indicator Kriging has no special requirements for regionalized variable to meet second-order stationary or normal distribution, so it has been widely used in many fields [6,7]. Due to space limitations, here we only take prediction of conditional probability of different sediment types as an example to briefly introduce its basic theory, and more details can be found in the references [5, 8].

Suppose a number of sediment samples have been obtained from a given sea area and their mass fractions of different grains are also determined in the laboratory. If Z refers to the combining conditions of a certain sediment type, that is the composition of mass fractions of sand, slit and clay, the indicator function at site x can be defined by eq. (1)

$$I(x,\ Z) = \begin{cases} 1 & z(x) \in Z \\ 0 & z(x) \notin Z \end{cases} \tag{1}$$

where $z(x)$ refers to the actual composition of the mass fractions of sand, slit and clay at site $x$, and $z(x) \in Z$ signifies the grain composition meets the requirements of grain composition of a specified type of sediment at site $x$. Under the given condition of Z, the probability of $z(x) \in Z$ is:

$$F(z) = \mathrm{Prob}[I(x,\ Z) = 1] \tag{2}$$

and the expectation of indicator function $I(x, Z)$ is:

$$E\{I(x,\ Z)\} = 1 \times F(z) + 0 \times [1 - F(z)] = F(z) \tag{3}$$

Eq. (3) makes clear on, at site $x$, the probability of the sediment belonging to the specified sediment type is equal to the average of indicator variables. If $I(x+h, Z)$ and $I(x, Z)$ are the random indicator variables separated by a vector $h$, the indicator semi-variance $\gamma_I(h, Z)$ can be defined as:

$$\gamma_I(h,Z) = \frac{1}{2} E\{[I(x+h,Z) - I(x,Z)]^2\} \tag{4}$$

According to eq. (4), a scatter plot of indicator semi-variance $\gamma_I(h, Z)$ against lag $h$

can be obtained and the plot can also be fitted with a theoretical semi-variogram model, such as spherical model, exponential model and exponential model. Accordingly, the ordinary Kriging method can be further used to predict $I(x,Z)$, the occurring probability of specified sediment type, at non-sampling site $x$ and its computation is:

$$F^*(Z) = \sum_{k=1}^{n} \lambda_k I(x_k) \tag{5}$$

where $F^*(Z)$ is the predicted value of occurring probability of the specified sediment type; $n$ is the number of samples in a neighborhood of the predicted site $x$; $x_k$ is the $k$th sample site in the neighborhood and $\lambda_k$ is its weight, which is determined by the indicator semi-variogram model under the condition of best linear unbiased estimation.

## 3 An application case of non- parametric mapping method

### 3.1 The study area and data

The study area covers the south of Lianyungang Harbor, north of Guanghe estuary, east of lower-water line and west of about -12 m isobath. In September 2005, a non-grid sampling scheme with the sampling transects almost perpendicular to the coastline was taken and a total of 106 surface sediment samples were collected. All the samples were located by GPS (Global Position System) and Fig. 1 showed their distribution map. The samples were analyzed in laboratory to obtain their mass fractions of different grains and then they were classified and nominated according to Shepard's classification system (see in Fig. 1). To facilitate the evaluation of mapping results, all samples were randomly divided into two groups. One group of 96 samples, called interpolation data set, were used for mapping, and the other group of 10 samples, named validation data set, were used for testing the accuracy of mapping results. The spatial distributions of the two data sets were also showed with different symbols in Fig. 1.

### 3.2 Results and analysis

#### 3.2.1 Calculations for indicator semi-variance of different types of sediment

According to Shepard's classification system, all the sediment samples can been classified into 6 types, namely sand, slity sand, sand-slit-clay, slit, sandy slit and clayey slit (Fig. 1). As such, for a specified type of sediment, each sample of the interpolation data set can be transformed into indicator variables of 1 and 0, and thus its indicator semi-variance can be calculated by using the geostatistics software GS+7.0. And then the scatter plot of semi-variance against lag $h$ can be drawn and fitted by a certain kind of theoretical model.

The scatter plots and theoretical models of six types of sediment are presented in Fig. 2.

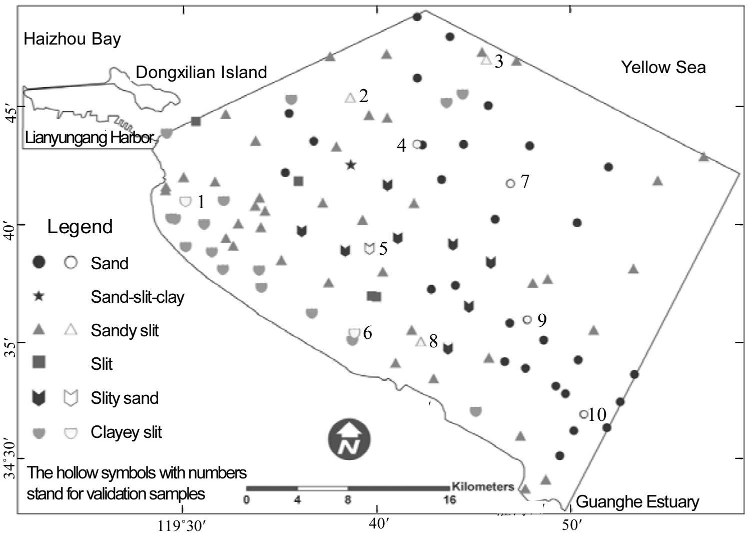


**Fig. 1　Sampling sites and their sediment types based on Shepard's classification system**

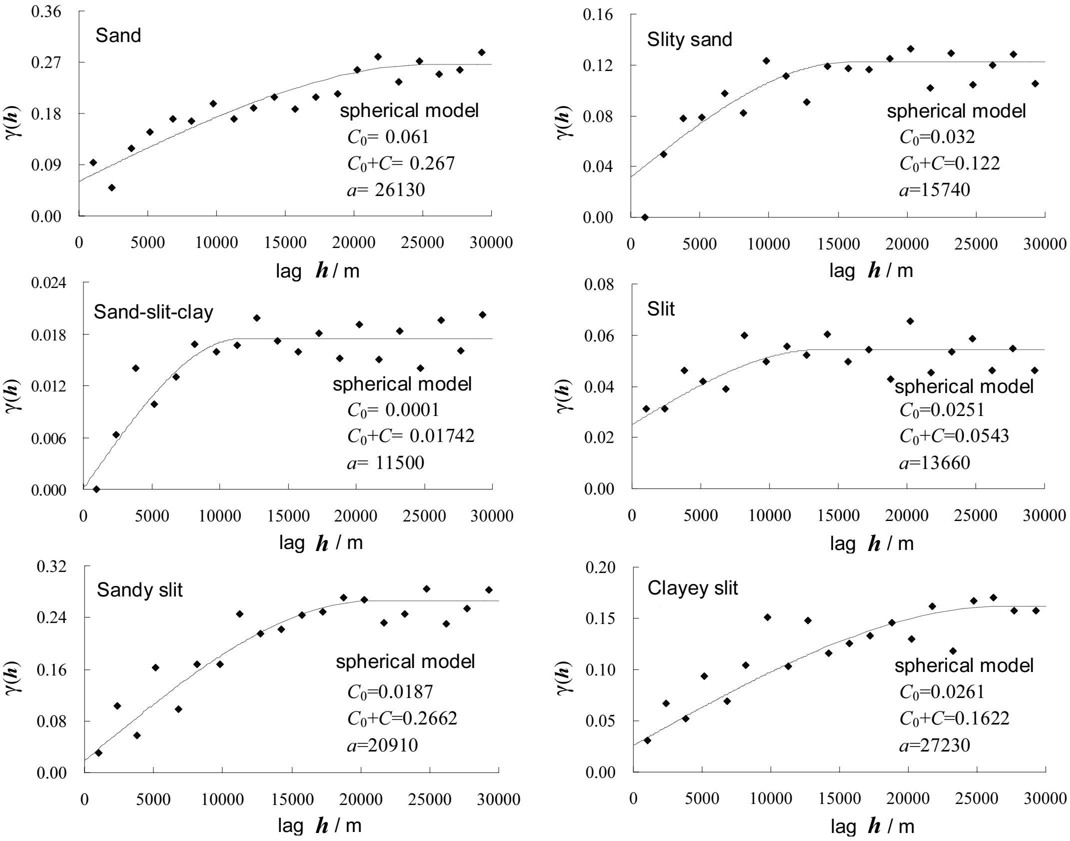


**Fig. 2　Semi-variograms of different types of sediment and their fitted models and parameters**

Fig. 2 shows that the indicator semi-variograms of six sediment types can be fitted with spherical models, indicating that all of six indicator semi-variograms have obvious structural characteristics. Therefore, Kriging method can further be used to predict indicator variable, namely occurring probability, of each sediment type over the whole study area.

### 3.2.2 Spatial prediction of occurring probability of different types of sediment

According to the fitted theoretical models, the occurring probability maps of different types of sediments are made by using Kriging interpolation method of GS+7.0 (Fig. 3).

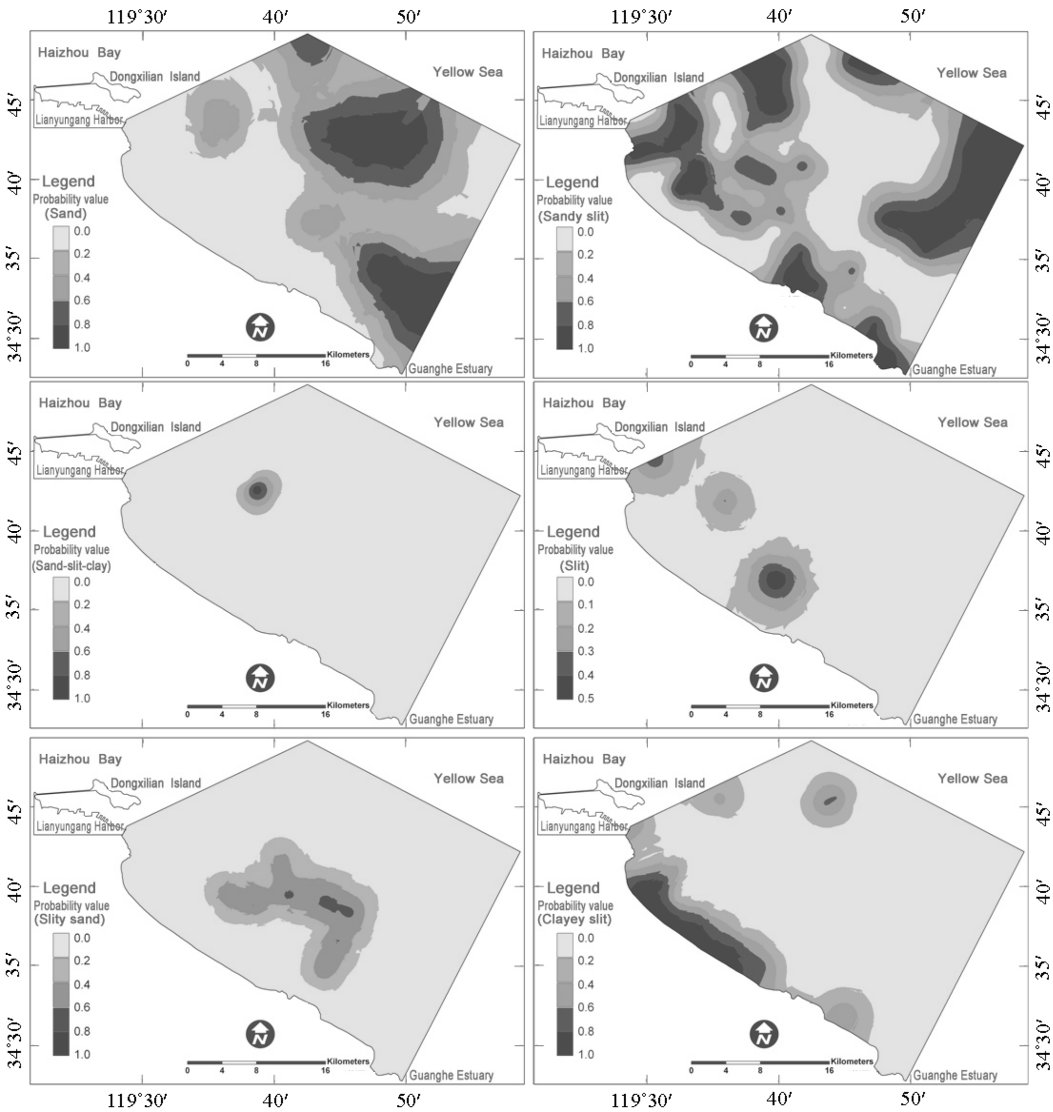


**Fig.3 Occurring probability distribution map of different types of sediment**

As can be seen from Fig. 3, the spatial distribution of occurring probability of six types of sediments is highly complementary and their higher values are spatially dislocated, which means different types of sediments dominate different areas respectively. Obviously, this phenomenon is very useful for generating sediment type map.

3.2.3 Generation of sediment type map and its uncertainty evaluation

With the occurring probability maps of different types of sediments, the sediment type map can be generated by using a kind of hardening method. The hardening method designates the sediment type with the highest occurring probability as the mapping type at a specified site, and this process can be implemented through map algebra functions of software ArcGIS. Fig.4 is the sediment type map generated by the hardening method. Comparing with Fig. 1, it can be found that the generated map not only presents clear boundaries among different types of sediments, but also preferably reflects their spatial distribution in the study area.

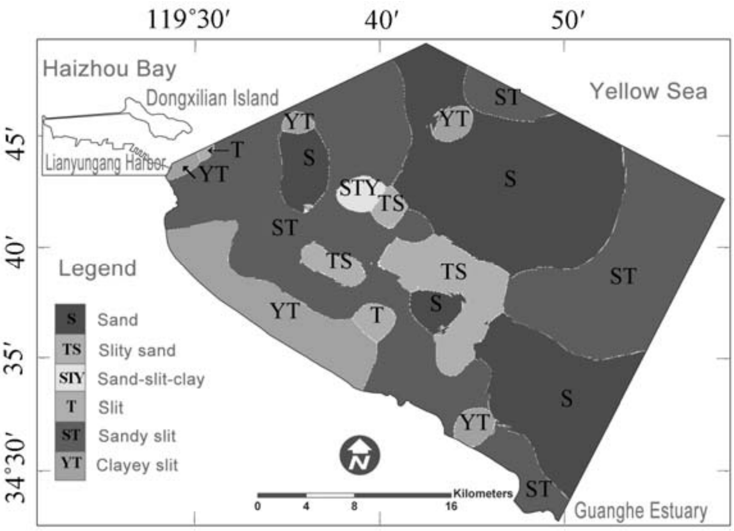


**Fig. 4 Sediment type map of southern sea area of Lianyungang**

As hardening process takes one type of sediment, it also ignores the occurring possibility of the other types, so hardening process also produces a kind of ignorance uncertainty. Ignorance uncertainty is usually relative to the gradual change of marine sediment, and when different types of sediment occur with similar probability values at a certain site, the ignorance uncertainty would be greater and the reliability of mapping result would also be lower. Therefore, we can generate an uncertainty map by calculating the ignorance uncertainty anywhere in the study area to evaluate the reliability of mapping results.

Usually, entropy is used to characterize the degree of ignorance uncertainty, because entropy can quantitatively reflects the degree of occurring probability tending to focus on a specified type of sediment. Entropy can be calculated as follows:

$$H_{ij} = -\frac{1}{\log_e n}\sum_{k=1}^{n}\left[\left(P_{ij}^{k}\Big/\sum_{l=1}^{n}P_{ij}^{l}\right)\log_e\left(P_{ij}^{k}\Big/\sum_{l=1}^{n}P_{ij}^{k}/P_{ij}^{l}\right)\right] \tag{6}$$

where $P_{ij}^{k}$ is the occurring probability of the $k$th type of sediment at the site of cell ($i$, $j$); $l$ is the number of known sediment types in the study area (in this paper, $l$ = 6); $H_{ij}$ , with its values range from 0 to 1, denotes entropy at the cell ($i$, $j$). When $H_{ij}$ = 0, cell ($i$, $j$) completely belongs to one of the six types of sediment and the hardening process does not produce any ignorance uncertainty; when $H_{ij}$ = 1, all the known types of sediment occur at cell ($i$, $j$) with equal probability, and putting cell ($i$, $j$) under any type of sediment would produce the greatest ignorance uncertainty. According to eq.(6), the entropy at each cell is calculated by using the map algebra functions of ArcGIS, and then the entropy distribution map, namely ignorance uncertainty map, is also obtained (Fig. 5).

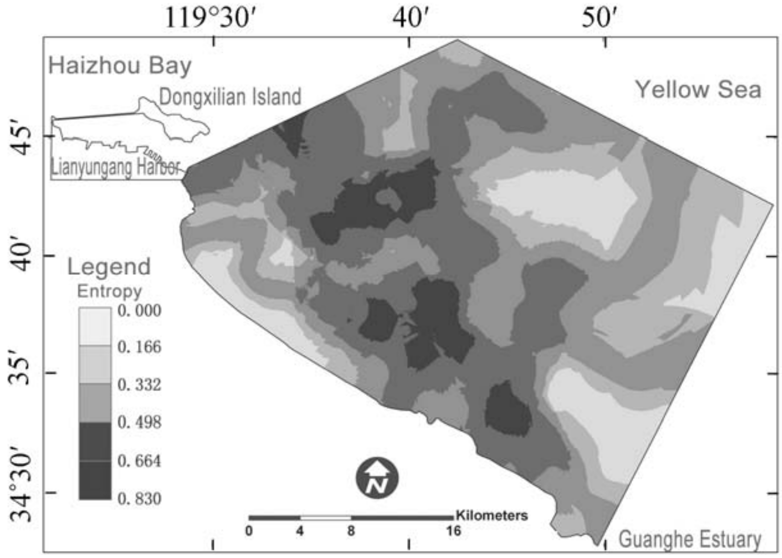


**Fig. 5 Spatial distribution of entropy for the sediment type map**

Fig. 5 shows that ignorance uncertainty in the southeast and northwest of the study area is relatively low, indicating their mapping reliability is high; while ignorance uncertainty in the central part shows patchy distribution with high values, indicating where the mapping results are somewhat unreliable. With Fig. 1, it can be found that high uncertainty is mainly distributed at the junction of various types of sediment samples, and low uncertainty mainly covers the areas where the same type of sediment samples are centralizedly distributed. Such phenomenon is in line with consequences of the general view, which indicates uncertainty map is worthy of applying to evaluate the reliability of the sediment type map.

### 3.2.4 Evaluation of mapping accuracy

For comparison and evaluation of mapping accuracy, Tab. 1 presents the actual and predicted types of 10 validation samples. Furthermore, their predicted results by the methods described in the reference [3, 4], namely the methods of IDW, Kriging and Voronoi diagram, are also listed in Tab. 1.

As shown in Tab. 1, the predicted results of non-parametric method are consistent with the actual results except the sample No. 5, while IDW and Kriging methods incorrectly predict four samples (sample No.2, 3, 4 and 5), which means the predicted accuracy of the non-parameter method is superior to that of IDW and Kriging methods. Although there is only one sample (sample No. 1) inconsistent with the actual results, predictions of Voronoi diagram method are usually sensitive to sample sites, and a slight offset of any sample site will lead to a big difference of mapping results, so its robustness is not good enough. In contrast, the non-parametric method has relatively low sensitivity to sample sites because its predictions are established on structural variation of indicator variables, therefore, its applicability is broader than that of Voronoi diagram method.

**Tab. 1 Comparison between the predicted types and actual types of the validation sediment samples**

| Sample ID | Actual type | Predicted type by non-parametric method | Predicted type by IDW method | Predicted type by Kriging method | Predicted type by Voronoi diagram method |
|---|---|---|---|---|---|
| 1 | clayey slit | clayey slit | clayey slit | clayey slit | sandy slit |
| 2 | sandy slit | sandy slit | slity sand | slity sand | sandy slit |
| 3 | sandy slit | sandy slit | slity sand | slity sand | sandy slit |
| 4 | sand | sand | sandy slit | slity sand | sand |
| 5 | slity sand | sandy slit | sandy slit | sandy slit | slity sand |
| 6 | clayey slit | clayey slit | clayey slit | clayey slit | clayey slit |
| 7 | sand | sand | sand | sand | sand |
| 8 | sandy slit | sandy slit | sandy slit | sandy slit | sandy slit |
| 9 | sand | sand | sand | sand | sand |
| 10 | sand | sand | sand | sand | sand |

## 4 Conclusions

The method for generating sediment type map based on a non-parametric indicator Kriging method can effectively avoid the subjectivity in the mapping process, and has no special requirements for sample data to meet second-order stationary or normal distribution. Furthermore, the method can also provided useful information on the quantitative evaluation of mapping uncertainty, which will help to enhance effective application of sediment type map. Through the application in the southern sea area of Lianyungang, the non-parameter method has obtained more convincing mapping results than the traditional methods such as IDW, Kriging and Voronoi diagram under the same condition. So the non-parameter method proposed in this paper is applicable with great utilization value.

It should be pointed out that the distribution of sediment types is usually effected by underwater topography and hydrodynamic condition, while in the methodological system introduced in the paper, the two important factors have not been considered, so the application of the non-parameter method in areas with complex topography and hydrodynamic conditions needs to be further tested, and it is also worthy of further discussing how to bring the relationship between spatial distribution of sediment and changes of underwater topography and hydrodynamic conditions into the proposed non-parameter methodological system.

## Acknowledgements

This research was financially supported by Natural Science Fund for colleges and universities in Jiangsu Province（No. 07KJD170012）and Natural Science Fund of Huaihai Institute of Technology (No. Z2008009).

## References

[1] ‘908’ Project Office of State Oceanic Administration. Technical manual of marine sediment investigation [M]. Beijing: Ocean Press, 2006: 33.

[2] Liu X Q. Map of the newest bottom-material types of china offshore continental shelf [J]. Marine Geology & Quaternary Geology, 1992,12(4):11 - 20.

[3] Zhang L H, Cui G S, Zhang J J, et al. A visualization method and its applications for information overlapping of submarine’s bottom characteristic and terrain [J]. Science of Surveying and Mapping, 2007, 32(4):111 - 112,115.

[4] Yang K, Zhang Y Z. A method to generate types and distribution of surface sediment map based on raster overlay function of GIS [J]. Quaternary Sciences, 2007,27(5):889 - 895.

[5] Journel A G. Non2parametric estimation of spatial distribution [J]. Math Geol, 1983,15(3): 445 - 468.

[6] Liu R M, Wang X J, Zhang W. Regional Environment Risk Assessment and Probability Distribution of Topsoil PAHs in Tianjin Area [J]. Environmental Science, 2008,29(6): 1 719 - 1 723.

[7] Xu Y, Chen Y X, Wang J S, et al. Using indicator Kriging to analyze and evaluate spatial distributions of soil water and salt in field [J]. Advances in Water Science, 2006 ,17(4) :477 - 482.

[8] Hou J R. Theory and Practice for Indicator Kriging [J]. Geology and Prospecting, 1990,26 (3):28 - 38.

[9] Zhu A X. Measuring uncertainty in class assignment for natural resource maps under a similarity model [J]. Photogrammetric Engineering & Remote Sensing, 1997,(63):1 195 - 1 202.

# 一种近海底质类型图生成的非参数方法

刘付程 [1]，彭 俊 [2]，张存勇 [1]

（1. 淮海工学院 测绘工程学院，江苏 连云港 222005；

2. 华东师范大学 河口海岸学国家重点实验室，上海 200062）

**摘 要：**近海底质类型图在近海工程和经济活动中有着广泛的应用价值。针对传统制图方法中存在的问题，本文提出了一种基于非参数指示Kriging的底质类型图生成方法。该方法能够有效地规避制图过程中的主观性，且对取样数据的平稳性和统计分布没有特殊要求，并能对制图结果的不确定性进行定量评价。该方法在连云港南部海域的应用实践表明，在相同的条件下，该方法可获得比传统方法更为精确地的制图结果，具有一定的实用价值。

**关键词：**沉积物类型；非参数指示Kriging；不确定性；制图