Acta Oceanol. Sin., 2017, Vol. 36, No. 2, P. 111–121
DOI: 10.1007/s13131-017-0969-3
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Optimization of Shanghai marine environment monitoring sites by integrating spatial correlation and stratified heterogeneity

FAN Haimei[1], GAO Bingbo[2*], XU Ren[1], WANG Jinfeng[3]

[1] East China Sea Environmental Monitoring Center, State Oceanic Administration, Shanghai 201206, China

[2] Beijing Research Center for Information Technology in Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100037, China

[3] State Key Laboratory of Resources & Environmental Information System, Institute of Geographic Sciences and Nature Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Received 16 October 2015; accepted 28 March 2016

Abstract

The water quality grades of phosphate (PO_4-P) and dissolved inorganic nitrogen (DIN) are integrated by spatial partitioning to fit the global and local semi-variograms of these nutrients. Leave-one-out cross validation is used to determine the statistical inference method. To minimize absolute average errors and error mean squares, stratified Kriging (SK) interpolation is applied to DIN and ordinary Kriging (OK) interpolation is applied to PO_4-P. Ten percent of the sites is adjusted by considering their impact on the change in deviations in DIN and PO_4-P interpolation and the resultant effect on areas with different water quality grades. Thus, seven redundant historical sites are removed. Seven historical sites are distributed in areas with water quality poorer than Grade IV at the north and south branches of the Changjiang (Yangtze River) Estuary and at the coastal region north of the Hangzhou Bay. Numerous sites are installed in these regions. The contents of various elements in the waters are not remarkably changed, and the waters are mixed well. Seven sites that have been optimized and removed are set to water with quality Grades III and IV. Optimization and adjustment of unrestricted areas show that the optimized and adjusted sites are mainly distributed in regions where the water quality grade undergoes transition. Therefore, key sites for adjustment and optimization are located at the boundaries of areas with different water quality grades and seawater.

Key words: area of water quality grade, stratified Kriging (SK), leave-one-out cross validation method, spatial simulated annealing method, monitoring sites optimization

Citation: Fan Haimei, Gao Bingbo, Xu Ren, Wang Jinfeng. 2017. Optimization of Shanghai marine environment monitoring sites by integrating spatial correlation and stratified heterogeneity. Acta Oceanologica Sinica, 36(2): 111–121, doi: 10.1007/s13131-017-0969-3

1 Introduction

Many countries have investigated the environmental status of estuaries, bays, and offshore marine areas. The US made guidelines for monitoring and assessing environment of all coastal waters (US EPA, 2010). And many marine environment monitoring programs and monitoring conventions has been made, for the Baltic Sea, the Mediterranean Sea, the North Sea, the East Asian seas and so on (Karydis and Kitsiou, 2013). The result told us that eutrophication is a serious marine environmental problem all over the world (Lindkvist et al., 2013). Baltic Marine Environment Protection Commission (Helsinki Commission) pointed out that because of nutrient pollution, most areas of the Baltic Sea are facing eutrophication problem (HELCOM, 2010). Analyzing results from the data collected between 2001 and 2006 showed that 161 of the total 172 coastal areas were affected by eutrophication. The eutrophication of the Black Sea was once much more serious than the Baltic Sea (Borysova et al., 2005), however it was found that the condition has been improving, the nitrogen and phosphate exported by the river has decreased by 20% between 1970 and 2000 (Karydis and Kitsiou, 2013).Thus continuous monitoring on the marine environment is necessary to reveal the dynamic variation of environmental quality and providing basis for environmental policy making (Karydis and Kitsiou, 2012; Wang et al., 2016).

Shanghai is located at the coastal confluence of the Changjiang (Yangtze River) Estuary and East China Sea, facing the East China Sea, north of the sea area of Jiangsu, and south to the sea area of Zhejiang. The Shanghai’s sea area has emerged as one of the most severely polluted sea areas in China because it is affected by the rivers flowing into the sea and land-sourced pollutants. In particular, changes in the material input and distribution of nutrients heavily influence this area (Wang et al., 2002; Shi et al., 2003; Zhou et al., 2006; Li et al., 2007). Chen and Chen (2003) showed the trend of water quality in the Shanghai’s sea area and

Foundation item: The National Natural Science Fundation of China under contract Nos 41376190, 41271404, 41531179, 41421001 and 41601425; the Open Funds of the Key Laboratory of Integrated Monitoring and Applied Technologies for Marin Harmful Algal Blooms, SOA under contract No. MATHA201120204; the Scientific Research Project of Shanghai Marine Bureau under contract No. HuHaiKe2016-05; the Ocean Public Welfare Scientific Research Project, State Oceanic Administration of the People's Republic of China under contract Nos 201305027 and 201505008.
*Corresponding author, E-mail: gaobb@lreis.ac.cn

demonstrated that the water quality near the river mouth bar has significantly deteriorated by nearly four times the growth of its nitrate content over the last 20 years. The China Oceanic Environment Quality Bulletin 2000 showed that the water quality at the Xuliujing section of the Changjiang River may be classified as poorer than Grade IV①. Over the last 20 years, the water quality in the Changjiang Estuary has continuously deteriorated; in particular, the water quality has decreased by at least one grade every five years. The deterioration of water quality in this area continues to increase. Eutrophication of water has progressed as follows: no eutrophication, 1980s; mild eutrophication, end of the 1980s and the beginning of 1990s; medium eutrophication, mid and late 1990s; and medium or severe eutrophication, 21st century (Quan et al., 2005; Han et al., 2003). The annual increase in dissolved inorganic nitrogen (DIN), nitrate, and phosphate (PO_4-P) discharges to the Changjiang River is consistent with the annual increase in their contents in the Shanghai's sea area. The annual growth rate of DIN in both areas is within the range of 3.3%–4.8%, while that of PO_4-P is within the range of 5.6%–6.2% (Fan et al., 2015). The community structure of phytoplankton in the Shanghai's sea area has constantly changed from 1984 to 2010, and percentages of dinoflagellate and diatom have changed in a manner consistent with the change in the percentages of nutrients in the Changjiang River. The percentages of DIN and PO_4-P, as well as discharges of silicate, have declined (Huang et al., 1986; Yang et al., 1989; Fang, 2008; Li, 2009).

The marine environment protection plans of domestic coastal provinces and cities in China indicate that polluted areas are management targets. One of the coastal cities or provinces expressed in its marine environment protection plan for 2011–2015 show that, "Marine environmental pollution will be under full and effective control, and environment quality will be significantly improved. The area of sea water with the sea water quality reaching or exceeding Grade II is estimated to increase by 5% in 2015. Environmental quality management targets defined in the hierarchy control program will be basically achieved". The change in the percentages of areas with different sea water qualities is reflected in the marine environment protection plan and in the annual marine environment quality communiqués of state/sea areas/provinces and cities. The percentage of areas with Grades I and II water quality is expected to increase, and areas with at least Grade IV water quality is expected to decrease. In recent years, the Shanghai Marine Environment Quality Communiqué has shown that sea areas with water quality poorer than Grade IV remain very high and account for at least 70% of such areas.

The existing monitoring sites must be optimized on the basis of nutrients (DIN and PO_4-P) to improve the efficiency of water quality monitoring and meet the needs of environmental management. An accurate mechanism or model is difficult to achieve in a marine environment because of the complex distribution of polluted sources and transport of pollutants. Therefore, the statistical method is generally adopted to optimize sites and perform statistical inferences (Shen and Wu, 2013; Sheikhy Narany et al., 2014). Based on regionalized variable theory and considering the variogram as a basic tool, the Kriging method makes use of spatial correlation for mapping of structural phenomenon with certain randomness (Matheron, 1963, 1967). Error variance of Kriging interpolation for target variable can be used to optimize the sampling patterns effectively and further improve the accuracy of mapping. It is widely used in marine, soil, and other environmental survey and monitoring (Shen and Wu, 2013; van Groenigen et al., 2000; Zimmerman, 2006). Kriging interpolation is based on spatial stationary hypothesis. However, spatial heterogeneity is one of the main characteristics of the study objects, so spatial sampling and statistical inference must be given careful attention (Wang et al., 2010). The nutrients in the Shanghai marine show obvious stratified heterogeneity which do not satisfy the spatial stationary requirements of Kriging and cannot be removed by fitting a spatially continuous surface (Gao et al., 2015). Goovaerts (1997) introduced stratified Kriging (SK, called Kriging with strata in that book) for regionalized variables with hierarchy and heterogeneity to produce best linear unbiased estimator. However, the spatial sampling optimization for mapping of the regionalized variables with stratified heterogeneity has not been resolved. Also, in the marine environment monitoring, the precision of estimated area of water quality grade is of much importance and much attention should be paid to.

This paper firstly implemented a comprehensive stratification for the study area and fitted global and local semi-variograms of nutrient elements, secondly selected the best statistical inference method for nutrient elements by cross validation. And, finally optimized current monitoring sites using bidirectional optimization method, that's to remove redundant sites according to the variation of areas of water quality grades and mean interpolation error, and then add new sites to right places to minimize the interpolation errors.

2 Data and methods

2.1 *Study area*

The 6 300-km long Changjiang River is the largest river in China, with a basin of 1.8 million km^2. The annual quantity of water flowing into the sea exceeds 960 billion m^3. Thus, the Changjiang River is the third largest river in the world. The Changjiang Estuary has become a region with a highly developed economy and very dense population. Hence, the estuary is considerably affected by human activities and must with stand tremendous pressure from resource usage and the environment. The Shanghai's sea area is located at the core of the Changjiang Estuary. In this study, data of the water quality in the Shanghai's sea area in August 2011 were collected in 70 monitoring sites. These data were derived from The East China Sea Environmental Monitoring Center, State Oceanic Administration, China. The study area is located at 30°30′–32°00′N, 121°00′–123°00′E (Fig. 1), where the big pints are monitoring site, and small points all over the water area are gotten by discretizing the study area. The surface content data of the water quality parameters, such as DIN and PO_4-Pwere used. The DIN included ammonia nitrogen, nitrate, and nitrite.

Samples were collected in the surface water (approximately 0.5 m deep). The samples were kept in well-sealed sampling bottles for cryopreservation. The samples were immediately analyzed after they were thawed in the laboratory. Analysis was based on *The Specifications of the National Standard of the People's Republic of China for Marine Monitoring*: *Part IV Sea*

①According to the National Standard of the People's Republic of China—Sea Water Quality Standards (GB3097-1997). Beijing: Standards Press of China

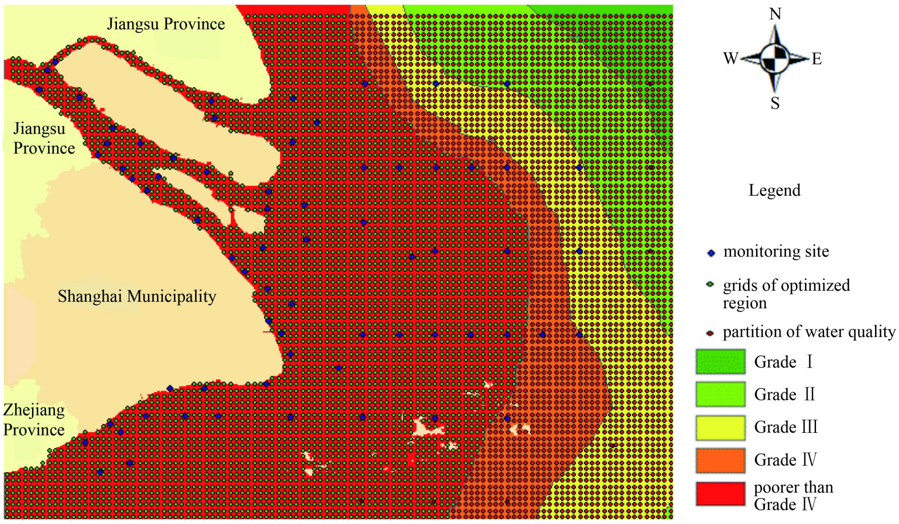

Fig. 1. Map of Shanghai's sea area. The big points are water quality monitoring sites, and small point are grids gotten by discretizing the study area.

Water Analysis (GB 17378.4-2007). The cadmium-copper column reduction method, hydrochloride naphthalene ethylenediamine spectrophotometry method, and indophenol blue spectrophotometry method were used to analyze ammonia nitrogen, nitrate, and nitrite, respectively. Phosphorus molybdenum blue spectrophotometry method was used to analyze PO_4-P.

2.2 *Ordinary Kriging (OK) and SK interpolation*

OK is a statistical method that uses spatial correlation for spatial interpolation. This process expresses the variable of the point to be estimated as the linear weight of peripheral points to solve the weighted coefficients on the basis of space stationary hypothesis, with unbiased and optimal conditions considered as limitations (Isaaks and Srivastava, 1989).

In OK, the observation values of the sample points are assumed to be (z_1, z_2, ..., z_n), while the actual value of the point to be estimated is z_0. The linear combination of the observation values of sample points is used to estimate the values of un-sampled points, as follows:

$$\hat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i. \tag{1}$$

In OK, the observation value of each position is a realization of random variable. Suppose expected values of the random variables in the study area are constants, i.e., they do not change with spatial position.

A spatial semi-variogram is used to reflect the spatial variation of the target variable. According to the stationary hypothesis, the observation values of multiple sample points can be considered as multiple observation values of a random variable. Therefore, a pair of points, with a distance of h between them, is used to calculate the actual semi-variogram:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left[z(u_{\mathrm{a}}) - z(u_{\mathrm{a}} + h) \right], \tag{2}$$

where $N(h)$ is the number of points pairs with the distance h. u_{a} is the start point of a point pair, and $u_{\mathrm{a}}+h$ is the end point of the point pair.

With unbiased and optimal conditions as limitations, the estimated weights without sampling might be to solve linear equations of the semi-variogram.

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \gamma(Z_1, Z_j) - \mu = \gamma(Z_0, Z_1), \\ \sum_{j=1}^{n} \lambda_j \gamma(Z_2, Z_j) - \mu = \gamma(Z_0, Z_2), \\ \quad \vdots \\ \sum_{j=1}^{n} \lambda_j \gamma(Z_n, Z_j) - \mu = \gamma(Z_0, Z_n), \\ \sum_{i=1}^{n} \lambda_i = 1, \end{cases} \tag{3}$$

where $\gamma(Z_1, Z_j)$ refers to the semi-variogram value between sample points 1 and j, $\gamma(Z_0, Z_1)$ refers to the semi-variogram value between the point to be estimated and sample point 1, μ refers to the Lagrange multiplier, and λ_i refers to the weighted coefficient in Eq. (1). λ_i is substituted into Eq. (1) to obtain the estimated value. The estimated error variance can be calculated through Eq. (4):

$$\delta_E^2 = 2\sum_{i=1}^{n} \lambda_i \gamma(Z_0, Z_i) - \gamma(Z_0, Z_0) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \gamma(Z_i, Z_j). \tag{4}$$

From Eq. (4), the estimated error of the estimated point is only related to the semi-variogram and the position of sample point, not to the value of the sample point. Therefore, the estimated error can be used to optimize the sample points.

SK is used for regionalized variables with spatially stratified heterogeneity based on OK. It proceeds in following three steps (Goovaerts, 1997): (1) dived the study area into continuous strata to minimize the difference within stratum and maximize the difference among strata; (2) fit the semi-variogram of each stratum using sampling data belong to it; and (3) predict the values of un-sampled points with nearby sampling data of the same stratum using OK.

2.3 *Spatial simulated annealing*

Spatial simulated annealing originates from the simulation of annealing in thermodynamics. An approximate optimal solution is obtained within polynomial time by slowly reducing the temperature and continuously decreasing the value of objective function (van Groenigen et al., 1997, 1999). Sites can be optimized according to spatial simulated annealing through the following steps:

(1) Design the objective function $f(s)$, for sampling optimization, set the sample sizen, initial temperature T_0, and cooling rate α.

(2) Generate the initial sampling plan s_0 using the simple and random or uniform space sampling method. Set $s=s_0$ and calculate the value of $f(s_0)$. Set current temperature $T=T_0$.

(3) Select a sampling point from the current sampling plans, and search a new point with a random angle and distance around the selected sampling point to replace it, thereby generate a new sampling plan s_1. Calculate the corresponding objective function $f(s_1)$ optimized target function. The following formula is used to calculate the acceptance probability of the new solution s_1:

$$p_i = \begin{cases} 1 & f(s_1) \leqslant f(s), \\ \operatorname{Exp}\left(\dfrac{f(s) - f(s_1)}{T}\right) & f(s_1) > f(s). \end{cases} \tag{5}$$

A random number (rand) is generated in 0–1. If rand<p_i, s_1 is accepted and set $s=s_1$. Otherwise, s_1 is rejected.

(4) Judge whether to stop the iteration. If the stopping criterion has not been satisfied, decrease the temperature by setting $T=T\alpha$ and go to Step (3) to continue the iteration. Otherwise, stop the iteration outputs.

2.4 *Monitoring sites optimization method*

The aim of monitoring sites optimization is to increase the accuracy of DIN and PO_4-P interpolation mapping and the areas calculation of different environmental quality grades. The optimization method consists of the following steps:

(1) Acquire spatial heterogeneity and correlation parameters. Analyze the spatial correlations and heterogeneity of DIN and PO_4-P. If the spatially stratified heterogeneity is found, the study area is stratified and the semi-variograms of all strata are fit.

(2) Select a proper inference method. Use leave-one-out cross validation to compare different inference methods to select the one which produce smallest error to be used in sites optimization.

(3) Removed redundancy sites in sequence. DIN and PO_4-P interpolation errors and changes in the area of some water quality resulting from reduction of sample points are used to evaluate the redundancy of sample information. In each round, the redundant site to be removed is determined as following:

(a) Order the sites, and select the first site to be left out, i.e., i=1.

(b) Employ the rest sites to interpolate the DIN contents in all points in the study are using the method selected from Step (2), calculate the mean interpolation error variance, and compare changes in mean interpolation error variances before and after reduction.

(c) Employ the rest sites to interpolate the PO_4-P contents in all points in the study area using the method selected from Step (2), calculate the mean interpolation error variance, and compare changes in mean interpolation error variance before and after reduction.

(d) Use the interpolation result of DIN and PO_4-P contents of above two steps to classify the water quality, and compare changes in areas with different water qualities before and after reduction.

(e) Judge whether all sites have been left out once. If yes, go to Step (f); if not, select the next site to be left out, and go to Step (b).

(f) Determine the site to be removed in this round. The site to be removed is the one that cause the smallest change of mean interpolation error variance and water quality areas after removal; The removing iteration ends when the number of removed sites reaches the limit of adjustable sites, or reaches the largest change in DIN and PO-P interpolation errors and areas of water qualities set beforehand.

(4) Optimize new sites. The mean error variances of DIN and PO_4-P in each position in the study area are used as optimization objectivity, as shown in Eq. (6):

$$O = \frac{1}{N}\sum_{i=1}^{N}\delta^2_{\mathrm{DIN}i} + 100 \times \frac{1}{N}\sum_{i=1}^{N}\delta^2_{\mathrm{PO_4}i}, \tag{6}$$

where N refers to the number of spatial point location in the study area, $\delta^2_{\mathrm{DIN}i}$ refers to the interpolation error variance of DIN in i spatial point, and $\delta^2_{\mathrm{PO_4}i}$ refers to the interpolation error variance of PO_4-P in the i spatial point. Remaining sites are used as fixed sites and spatial simulated annealing is used to search the optimal locations for new sites. The optimization method was realized by coding with R language (V 2.15.1) and the gstat package (1.0) (Pebesma, 2004).

3 Results

3.1 *Spatial stratification*

Water quality in a location is determined by the index with poorest grade. The DIN and PO_4-P were used to calculate the water quality which inturn was used to stratify the study area (Fig. 2). Areas with water quality poorer than Grade IV were designated as Stratum 1and areas with water qualities of Grades I, II, III, and IV were combined to form Stratum 2. The Hangzhou Bay has a high DIN and PO_4-P concentration, similar to the DIN and PO_4-P value at the Changjiang River mouth, but far away from it. This character might affect the accuracy of semi-variogram if all sampling data were adopted without discrimination. To void this, the Hangzhou Bay was independently divided into a subzone of Stratum 1, called Region 2. The remaining area with water quality poorer than Grade IV is designated as another subzone of Stratum 1, called Region 1 (Fig. 3).

3.2 *Semi-variograms fitting*

The global and local semi-variograms of DIN and PO_4-P are fitted according to the spatial stratification result. For the global semi-variogram, all sites are included in the calculation. The semi-variogram of Region 1 only uses the sites within it. Because the sites in Region 2 were not enough and its spatial variation is similar to that of Region 1, it used semi-variogram of Region 1 as its semi-variogram. Thus the Stratum1 have only one uniform semi-variogram fitting using data in Region 1, and in calculating and analysis of rest part, Stratum 1 would not be divided into Region 1 and Region 2, but as one unit. For the semi-variogram in Stratum 2, all sample points are used for the fitting calculation.

The semi-variogram fitting results of DIN are listed in Table 1. The semi-variogram of Global, Stratum 1 and Stratum 2 are all of Gaussian type, but with different range, sill and nugget. Figure 4 listed their semi-variogram plots.

The PO_4-P contents are generally small numbers, and the fitting result can easily be affected by rounding. Therefore, the PO_4-

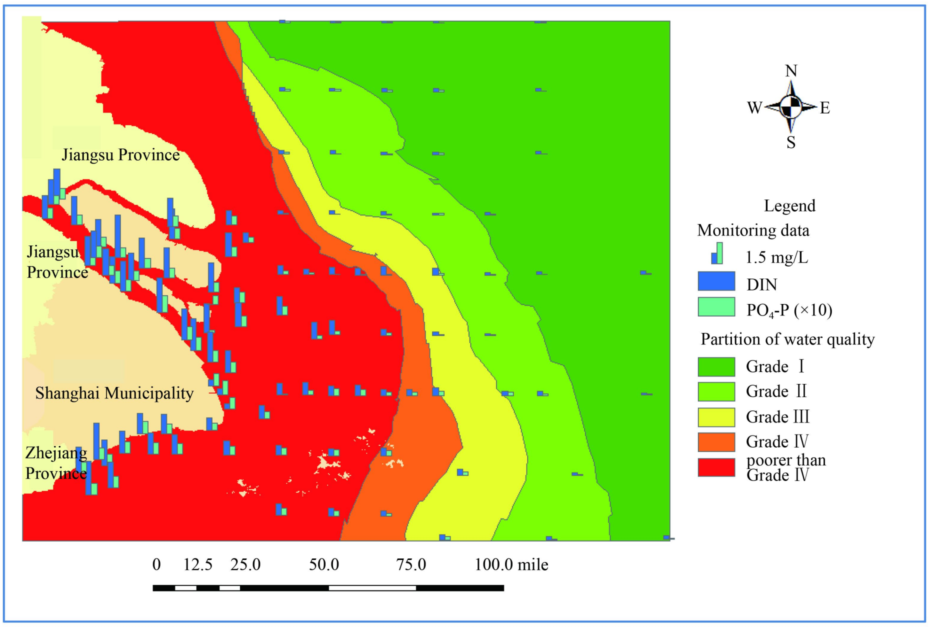

Fig. 2. Water quality map of the study area.

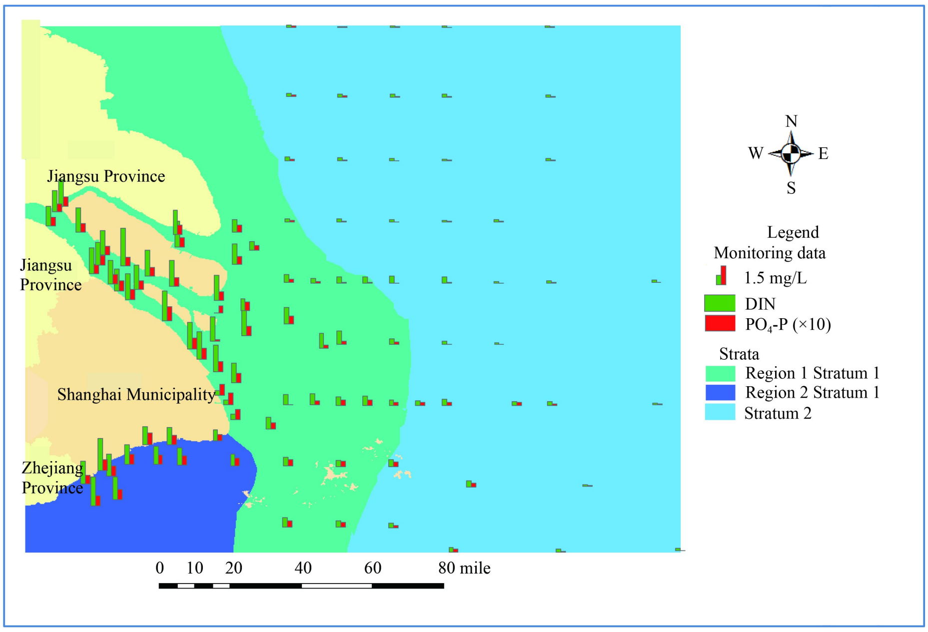

Fig. 3. Spatial stratification result. The study area were stratified into two stratum, and the Stratum 1 was further divided into two regions.

Table 1. Adjusted semi-variogram models for DIN concentration

	Type	Nugget (C_0)	Sill (C_0+C)	Range (a)
Global	Gaussian	0.220	1.210	190 945
Stratum 1	Gaussian	0.252	0.733	95 640
Stratum 2	Gaussian	0	1.508	25 013

P content is multiplied by 10 for the before calculation. The semi-variogram results of PO_4-P are listed in Table 2. Again, the semi-variogram of Global, Stratum 1 and Stratum 2 are all of Gaussian type, but with different range, sill and nugget. Figure 5 listed the corresponding semi-variogram plots.

3.3 *Selection of statistical inference method*

Based on analysis of the spatially stratified heterogeneities and correlations of DIN and PO_4-P, OK, SK, and inverse distance weighted (IDW) interpolation methods were selected preliminarily as candidate statistical inference methods. Cross validation method was used to select the best inference method. In the cross validation method, one site was left out and other sites were used to interpolation the DIN or PO_4-P of this missing site every time, until every site has been interpolation once. The absolute mean error and error variance of the interpolations are calculated and compared.

OK, SK, and IDW cross validation results of DIN are shown in

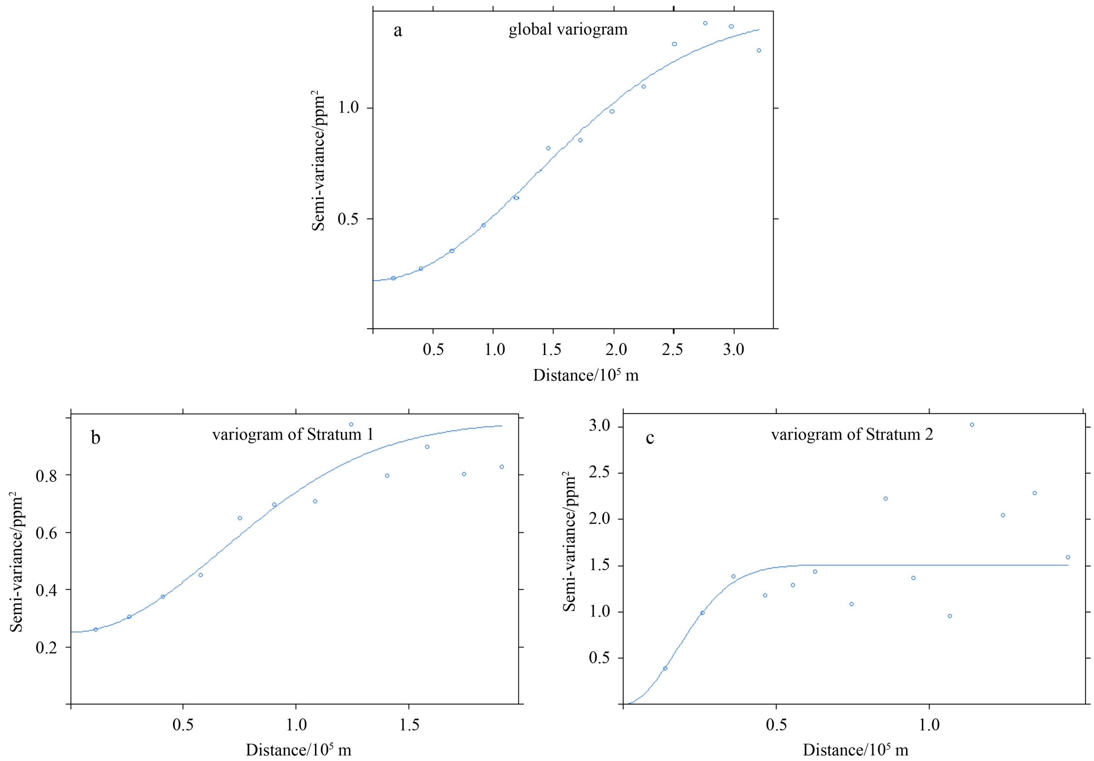

Fig. 4. Semi-variogram fitting result for the content of DIN. a. Variogram of the global, b. of Stratum 1, and c. of Stratum 2.

Table 2. Adjusted variogram models for PO_4-P concentration

	Type	Nugget (C_0)	Sill (C_0+C)	Range (a)
Global	Gaussian	0	0.206	166 914
Stratum 1	Gaussian	0	0.057	20 621
Stratum 2	Gaussian	0.075	0.773	53 140

Table 3. SK exhibits the smallest absolute average error and error variance; thus, it is selected for DIN. The error distribution is shown in Fig. 6. OK and SK interpolation errors gradually increase with improvements in water quality from Grade I to poorer than Grade IV. Water quality poorer than Grade IV shows

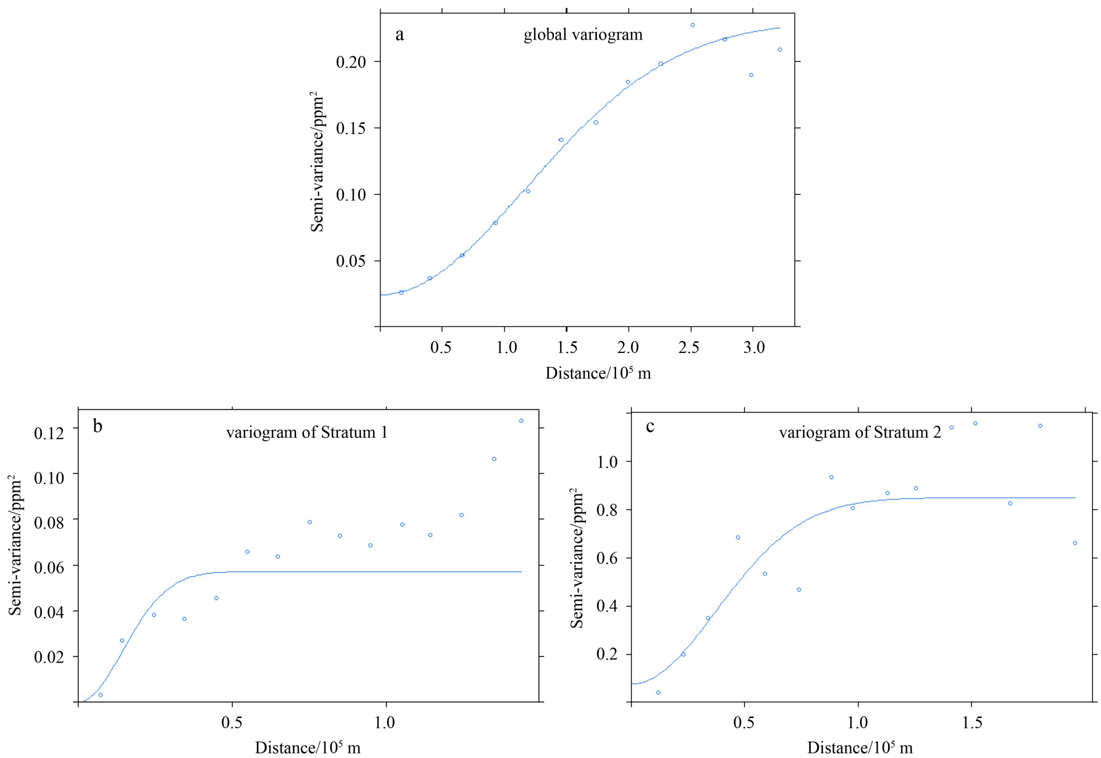

Fig. 5. Semi-variogram fitting result for the content of PO_4-P. a. Variogram of the global, b. of Stratum1, and c. of Stratum 2.

Table 3. DIN cross validation result

	OK	SK	IDW
Absolute average error	0.253 4	0.231 2	0.469 0
Error variance	0.125 9	0.117 0	0.277 8

the largest error. The absolute average errors of the various water qualities are as follows: Grade I is 0.054 1; Grade II is 0.053 6; Grade III is 0.098 8; Grade IV is 0.151 0; and poorer than Grade IV is 0.327 8.

The results of OK, SK, and IDW cross validation results of PO_4-P are shown in Table 4. OK reveals the smallest absolute average error and error variance; thus, it is selected for PO_4-P. The error distribution of the various methods is shown in Fig. 7. OK and SK interpolation errors gradually increase as water quality deteriorates from Grade I to poorer than Grade IV. Water quality poorer than Grade IV shows the highest error. The absolute average errors of the different water quality grades are as follows: Grade I is 0.004 56; Grade II is 0.005 48; Grade III is 0.008 76, Grade IV is 0.009 90; and poorer than Grade IV is 0.011 09.

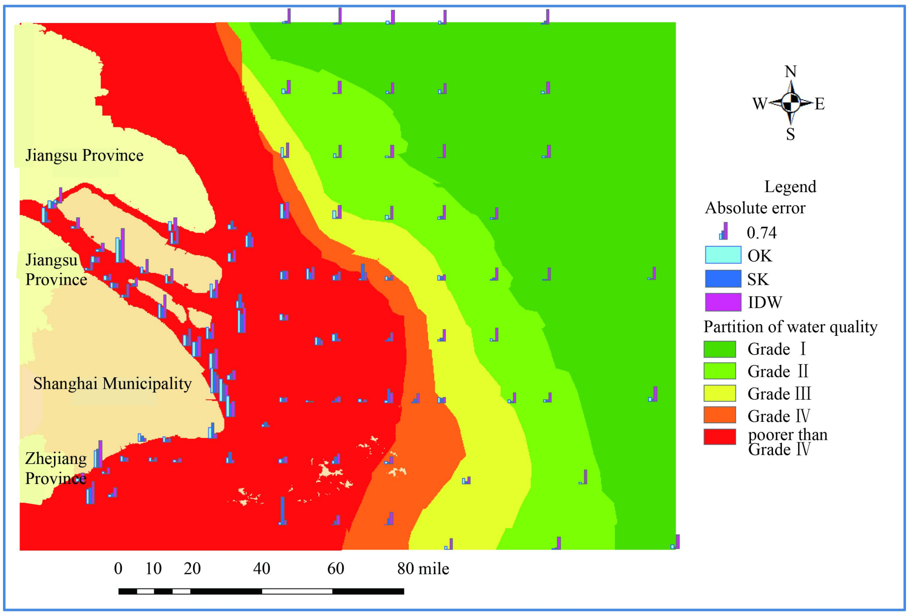

Fig. 6. Absolute error map of DIN cross validation.

Table 4. PO_4-P cross validation result

	OK	SK	IDW
Absolute average error	0.092 35	0.141 86	0.194 57
Error variance	0.016 77	0.051 58	0.047 43

3.4 *Reduction of redundant monitoring sites*

Seven redundant historical sites (10% of the 70 monitoring sites) will be removed. Each site is gradually reduced based on the site reduction method. The reduced sites are labeled by their codes (Fig. 8). Table 5 shows the interpolation errors obtained after gradual reduction of sites, gradual changes, and accumulated changes in the areas of various water quality grades. The area of each grid is approximately 4 km^2.

After A2D31YQ015S is removed, the corresponding mean interpolation deviations of DIN and PO_4-P change to 1.00×10^{-4} and 3.00×10^{-6}, respectively, resulting in slight changes in error. The area of poorer than Grade IV water quality decreases by 1 unit, whereas that of Grade IV increases by 1 unit. After A2D31YQ013S is removed, the corresponding mean interpolation deviations of DIN and PO_4-P change to 2.00×10^{-4} and 1.00×10^{-6}, respectively, which do not cause changes in the areas of various water quality grades. After A2D31YQ067S is removed, the corresponding mean interpolation deviations of DIN and PO_4-P change to 4.00×10^{-4} and 1.00×10^{-5}, respectively, which also do not cause changes in the areas of various water quality grades.

After A2D31YQ062S is removed, the corresponding mean interpolation deviations of DIN and PO_4-P change to 6.00×10^{-4} and 7.00×10^{-6}, respectively, which do not cause changes in the areas. After A2D31YQ030S is removed, the mean interpolation deviations of DIN and PO_4-P change to 1.50×10^{-3} and 3.00×10^{-6}, respectively. Areas of water qualities of Grades I and III are decreases by 1 unit, whereas those with water qualities of Grades II and IV increase by 1 unit. After A2D31YQ004S is removed, the mean corresponding interpolation deviations of DIN and PO_4-P change to 3.00×10^{-4} and 3.00×10^{-6}, respectively. Areas with water qualities of Grades I and IV increase by 1 unit, where as those with water quality poorer than Grade IV are reduced by 2 units. After A2D31YQ069S is removed, the corresponding mean interpolation deviations of DIN and PO_4-P change to 9.00×10^{-4} and 2.40×10^{-5}, respectively. Areas with water qualities of Grade I and poorer than Grade IV are reduced by 1 unit of area, whereas those with water qualities of Grades II and IV increase by 1 unit (Table 5).

Changes in the accumulated area are as follows: Grade I decreases by 1 unit; Grade II increases by 2 units; Grade III decreases by 1 unit; Grade IV increases by 4 units; and poorer than Grade IV decreases by 4 units (Table 5).

3.5 *Optimization of the monitoring sites*

3.5.1 *Under the restriction of water grades*

The sum of the interpolation errors of DIN and PO_4-P (observed values were multiplied by 10) in all positions is minim-

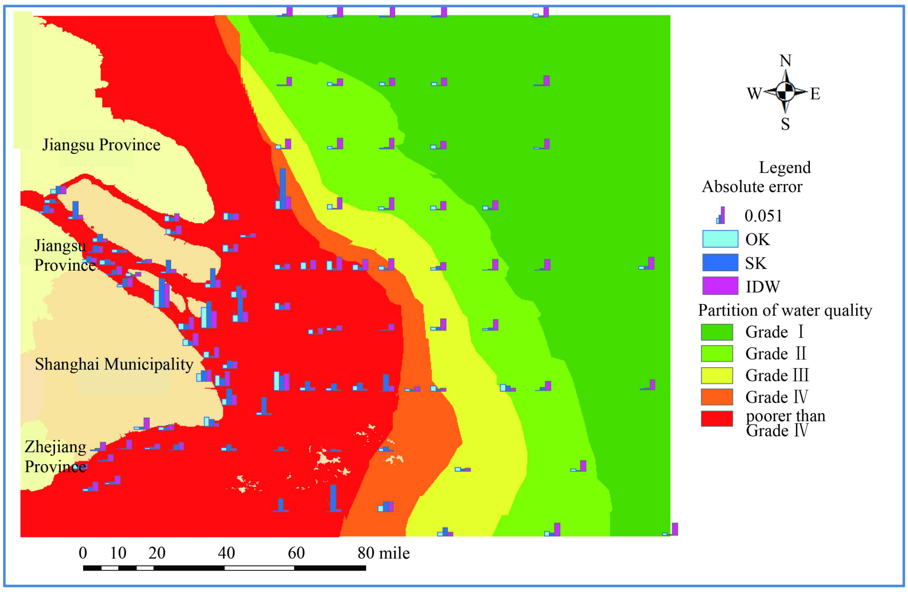

Fig. 7. Absolute error map of PO_4-P cross validation.

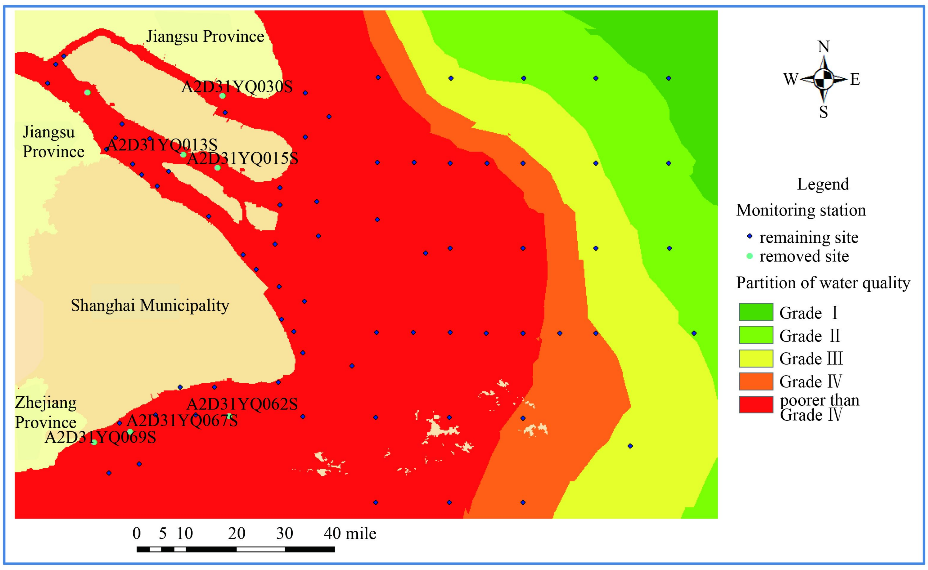

Fig. 8. Results map of sites reductions.

Table 5. Standard deviation of interpolation errors and (accumulated) changes in the areas of various water quality grades after site reduction

No. of sites	Standard deviation of DIN interpolation error	Standard deviation of PO_4-P interpolation error	Grade I	Grade II	Grade III	Grade IV	Poorer than Grade IV
All sites	0.404 8	0.016 177	669	678	741	640	4 123
A2D31YQ015S	0.404 9	0.016 180	0(0)	0(0)	0(0)	1(1)	–1(–1)
A2D31YQ013S	0.405 1	0.016 181	0(0)	0(0)	0(0)	0(1)	0(–1)
A2D31YQ067S	0.405 5	0.016 191	0(0)	0(0)	0(0)	0(1)	0(–1)
A2D31YQ062S	0.406 1	0.016 198	0(0)	0(0)	0(0)	0(1)	0(–1)
A2D31YQ030S	0.407 6	0.016 201	–1(–1)	1(1)	–1(–1)	1(2)	0(–1)
A2D31YQ004S	0.407 9	0.016 204	1(0)	0(1)	0(–1)	1(3)	–2(–3)
A2D31YQ069S	0.408 8	0.016 228	–1(–1)	1(2)	0(–1)	1(4)	–1(–4)

ized to add seven sits in the subzones of Grades III and IV water quality. SK interpolation is used to interpolate for DIN, whereas OK interpolation is used for PO_4-P. After optimization of the monitoring sites, the standard deviation of DIN interpolation errors of all positions is 0.399 82 and that of PO_4-P is 0.016 087.

Optimization and adjustment under the restriction of water grades show that the optimized and adjusted sites are mainly distributed in regions where the water quality grade undergoes transition, for example, two added sites near the transition of Grades II and III water quality, three near the transition of Grades III and IV, and one near the transition of Grade IV and the poorer than Grade IV (Fig. 9).

3.5.2 *Without any restriction*

The sum of the interpolation errors of DIN and PO_4-P (observed values were multiplied by 10) in all positions is minimized to optimize the selected points over the whole region without any limitation. SK interpolation is used to interpolate DIN, whereas OK is used for PO_4-P. After optimization of the monitoring sites, the standard deviation of DIN interpolation errors of all positions is 0.396 77 and that of PO_4-P is 0.016 071.

Optimization and adjustment of unrestricted areas also show that the optimized and adjusted sites are mainly distributed in regions where the water quality grade undergoes transition, for example, one added site near the transition of Grades I and II water quality, two near the transition of Grades II and III, and one near the transition of Grades III and IV, and so on (Fig. 10).

4 Discussion

DIN and PO_4-P in Shanghai Marine seriously exceed standard values. Thus, eutrophication of the estuary area must be monitored. Nutrients are the main substances of terrigenous input. Although local governments continue to enhance the partitioning of marine environment protection and seek environmental protection audits of sea areas, the marine environmental protection of sea areas is also under the jurisdiction of these governments. The proportional distribution of areas under different water quality grades is the first task in marine environmental protection. The increasing percentages of areas with water qualities of Grades I and II and the decreasing percentages of areas with water qualities of Grade IV and poorer than Grade IV are specific tokens of environmental protection performance. Thus, the total terrigenous pollutant to the sea is under control, and the discharge is up to the required standard.

Table 6. Position coordinates of optimized site in the subzones

North latitude/(°)	East longitude/(°)	Water partition
30.491 2	123.024 7	Grade III
31.591 6	122.898 6	Grade III
30.815 9	123.066 6	Grade IV
30.942 2	123.087 6	Grade III
31.122 6	123.045 9	Grade III
30.562 9	122.649 5	Grade IV
30.472 9	123.274 7	Grade III

Partitioning of the Shanghai's sea area on the basis of water quality grades or types of sediments (Chen et al., 2009) can reflect stable physical dynamic processes in this sea area. Under the effects of the runoff from the Changjiang River and co-oscillating tides of the open sea, the main reversing currents in the estuary are bounded by channel transport to the open sea (with obvious inequality of tides). The impact of tides from the open sea is significantly enhanced outside the estuary, where rotating flow is focused and a turbid zone with obvious stratification is formed. Sea area stratification (spatial heterogeneity) and semi-variogram fitting (space correlation) are used to solve various complex physical processes over the whole sea area; these methods can accurately reflecting the main water masses of the sea area and their distribution characteristics.

Optimization objectivity of monitoring sites is based on the monitoring purpose and program design. When focusing on the environmental characteristics of a comprehensive index, minimizing the error variance of the principle component of the comprehensive index is used as optimization target of the monitoring sites (Bing-Bo Gao et al., 2015). When focusing on the environment characteristics of nutrients (DIN and PO_4-P), the main elements exceeding the standard, and their distribution, the site impact on DIN and PO_4-P errors should be fairly low. Redundant sites with limited influence on the estimation of polluted area and the interpolation errors were removed, and several sites were added within the border of Grades III and IV or within the boundary of the whole sea area.

The characteristics of the sea area under different seasons (temporal change) and partitions (spatial change) can influence the selection of statistical inference methods and affect the op-

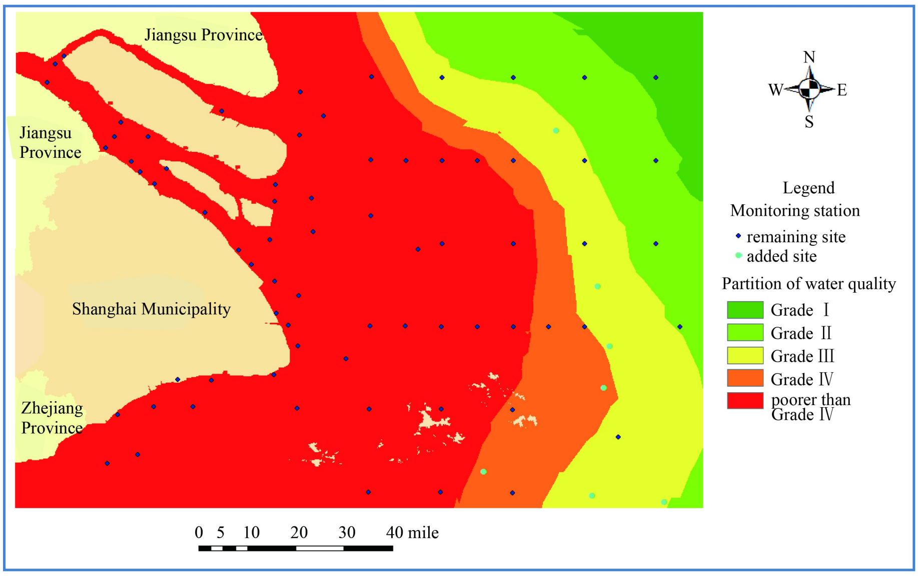

Fig. 9. Optimized site layout with the addition of seven sites in the subzones.

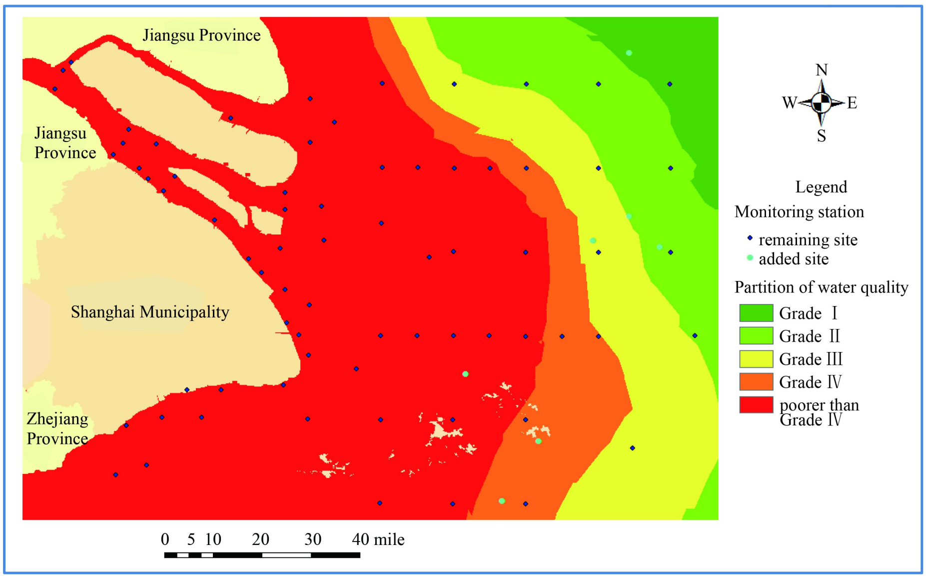

Fig. 10. Optimized site layout with the addition of seven sites in the whole region.

Table 7. Position coordinates of optimized site in the whole region

North latitude/(°)	East longitude/(°)	Water partition
31.266 7	123.213 9	Grade II
30.689 5	122.795 2	Grade IV
31.357 1	123.109 0	Grade II
31.285 0	122.982 9	Grade III
30.887 3	122.543 7	poorer than Grade IV
31.844 1	123.109 6	Grade I
30.508 8	122.670 5	Grade IV

timization results of the monitoring sites. However, for estuaries, bays, and other regions with more determinate characteristics of water mass, the statistical inference and its parameters of some key elements can be stable. For complex temporal and spatial variations in the sea area, researchers and managers must proceed from the whole situation to perform the necessary sea area partitioning and site adjustment based on targeted monitoring purpose and the distribution characteristics of the main physical water masses. This implement of the monitoring sites optimization must achieve multiplier effects.

5 Conclusions

The Shanghai Marine Environment Monitoring sites were optimized by integrating spatial correlation and stratified heterogeneity. The distribution of areas of different water quality grades in the Shanghai's sea area was generally determined by the poorest index. This process basically reflects the distribution characteristics of the water quality grades according to the nutrients. To calculate the areas under the different water quality grades according to the optimized site layout more accurately, spatial stratum was stratified by integrating the current DIN and PO_4-P status of sea water quality grades. The global and local semi-variograms of DIN and PO_4-P were fitted, and the statistical inference methods were selected through leave-one-out cross validation. SK interpolation was used for DIN, whereas OK interpolation was used for PO_4-P. During optimization of the monitoring sites, redundant monitoring sites that had the lowest impact on the areas of different water grade, and the interpolation mean errors of the DIN and PO_4-P were firstly removed. To adjust the historical sites by 10%, seven historical sites were removed. And then to improve the accuracy of DIN and PO_4-P interpolation mapping new sites were added optimally to right locations with or without restriction of water grades.

Using the original sites, the standard deviation of DIN interpolation errors of all positions was 0.404 8, whereas that of PO_4-P was 0.016 18. If the points in the strata with Grades III and IV are optimized, the DIN standard deviation of interpolation errors of all positions becomes 0.399 82 after optimization and that of PO_4-P becomes 0.016 087. If the points over the whole region are optimized without any limitation, the standard deviation of DIN interpolation error of all positions becomes 0.396 77 and that of PO_4-P becomes 0.016 071 after site optimization. Optimization method was applied to optimize the environmental monitoring sites in estuary areas; this process can improve the interpolation accuracy and computational accuracy of areas with different water quality grades.

References

Borysova O, Kondakov A, Palcari S, et al. 2005. Eutrophication in the Black Sea Region; Impact Assessment and Causal Chain Analysis. Kalmar: University of Kalmar

Chen Jiyu, Chen Shenliang. 2003. The changes of ecologic environment in Yangtze river estuary and some suggestions for estuary regulation. Water Resources and Hydropower Engineering (in Chinese), 34(1): 19–25

Chen Shenliang, Yan Shuzhuang, Li Yuzhong. 2009. Characteristics of surface sediment distribution in the Yangtze estuary and its adjacent waters. Resources and environment in the Yangtze Basin (in Chinese), 18(2): 152–156

Fan Haimei, Gao Bingbo, Yu Jiang, et al. 2015. The trend of nutrient contents in seawater around shanghai and an analysis on their correlation with fluxes via the Yangtze River. Shanghai Environmental Sciences (in Chinese), 34(1): 1–5, 25

Fang Qian. 2008. The study on sources of main chemical pollutants and fluxes flowing into the East China Sea in recent 30 years (in Chinese) [dissertation]. Qingdao: Ocean University of China

Gao Bingbo, Wang Jinfeng, Fan Haimei, et al. 2015. A stratified optimization method for a multivariate marine environmental monitoring network in the Yangtze River estuary and its adjacent sea. International Journal of Geographical Information Science, 29(8): 1332–1349

Goovaerts P. 1997. Geostatistics for Natural Resources Evaluation. New York: Oxford University Press

Han Xiurong, Wang Xiulin, Sun Xia, et al. 2003. Nutrient distribution and its relationship with occurrence of red tide in coastal area of East China Sea. Chinese Journal of Applied Ecology (in Chinese), 14(7): 1097–1101

HELCOM. 2010. Ecosystem Health of the Baltic Sea 2003–2007: HELCOM initial holistic assessment. In: Andersen J H, Korpinen S, Laamanen M, et al., eds. Baltic Sea Environmental Proceedings No. 122. Helsinki: Baltic Environmental Protection Commission, Helsinki Commission

Huang Shanggao, Yang Jiadong, Ji Weidong, et al. 1986. Spatial and temporal variation of reactive Si, N, P and their relationship in the Changjiang Estuary water. Taiwan Strait (in Chinese), 5(2): 114–123

Isaaks E H, Srivastava R M. 1989. Applied Geostatistics. New York: Oxford University Press

Karydis M, Kitsiou D. 2012. Eutrophication and environmental policy in the Mediterranean Sea: a review. Environmental Monitoring and Assessment, 184(8): 4931–4984

Karydis M, Kitsiou D. 2013. Marine water quality monitoring: a review. Marine Pollution Bulletin, 77(1–2): 23–36

Li Dan. 2009. The study on the hydro-chemical characteristics and the flux to the sea about the rivers in the East of China (in Chinese) [dissertation]. Shanghai: East China Normal University

Li Zheng, Shen Zhiliang, Zhou Shuqing, et al. 2007. Distributions and variations of phosphorus in the Changjiang estuary and its adjacent sea areas. Marine Science (in Chinese), 31(1): 28–36, 42

Lindkvist M, Gren I M, Elofsson K. 2013. A study of climate change and cost effective mitigation of the Baltic sea eutrophication. In: Singh B R, ed. Climate Change-Realities, Impacts Over Ice Cap, Sea Level and Risks. Rijeka, Croatia: InTech

Matheron G. 1963. Principles of geostatistics. Economic Geology, 58(8): 1246–1266

Matheron G. 1967. Kriging, or polynomial interpolation procedures. Canadian Mining and Metallurgical Bulletin, 60: 1041–1045

Pebesma E J. 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences, 30(7): 683–691

Quan Weimin, Shen Xinqiang, Han Jindi, et al. 2005. Analysis and assessment on eutrophication status and developing trend in Changjiang Estuary and adjacent sea. Marine Environmental Science (in Chinese), 24(3): 13–16

Sheikhy Narany T, Ramli M F, Aris A Z, et al. 2014. Spatial assessment of groundwater quality monitoring wells using indicator kriging and risk mapping, amol-babol plain, Iran. Water, 6(1): 68–85

Shen Yaqi, Wu Yanqing. 2013. Optimization of marine environmental monitoring sites in the Yangtze River estuary and its adjacent sea, China. Ocean & Coastal Management, 73: 92–100

Shi Xiaoyong, Wang Xiulin, Han Xiurong, et al. 2003. Nutrient distribution and its controlling mechanism in the adjacent area of Changjiang River Estuary. Chinese Journal of Applied Ecology (in Chinese), 14(7): 1086–1092

US EPA. 2015. National coastal condition assessment: site evaluation guidelines. https://www.epa.gov/sites/production/files/2016–09/documents/final_ncca15_seg_09–14-2016.pdf. [2015-04-20/2016-09-14]

van Groenigen J W, Pieters G, Stein A. 2000. Optimizing spatial sampling for multivariate contamination in urban areas. Environmetrics, 11(2): 227–244

van Groenigen J W, Siderius W, Stein A. 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma, 87(3–4): 239–259

van Groenigen J W, Stein A, Zuurbier R. 1997. Optimization of environmental sampling using interactive GIS. Soil Technology, 10(2): 83–97

Wang Jinfeng, Haining R, Cao Zhidong. 2010. Sample surveying to estimate the mean of a heterogeneous surface: reducing the error variance through zoning. International Journal of Geographical Information Science, 24(4): 523–543

Wang Juncheng, Wang Zhongqiu, Wang Yiming, et al. 2016. Current situation and trend of marine data buoy and monitoring network technology of China. Acta Oceanologica Sinica, 35(2): 1–10

Wang Baodong, Zhan Run, Zang Jiaye. 2002. Distributions and transportation of nutrients in Changjiang River Estuary and its adjacent sea areas. Haiyang Xuebao (in Chinese), 24(1): 53–58

Yang Xiaolan, Lin Yian, Zhang Jian. 1989. The characteristics of environmental hydrochemistry in the Changjiang Estuary and its adjacent area. Danghai Marine Science (in Chinese), 7(2): 60–65

Zhou Junli, Liu Zhengtao, Meng Wei, et al. 2006. The characteristics of nutrients distribution in the Yangtze River Estuary. Research of Environmental Sciences (in Chinese), 19(6): 139–144

Zimmerman D L. 2006. Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. Environmetrics, 17(6): 635–652