in the paper, but we have only analyzed a limited number of possible scenarios, and ideally we should take into account much more possibilities.

ACKNOWLEDGEMENTS

This research has been partially supported by grant BFM2001-3286. The referees are acknowledged for their helpful comments and suggestions.

REFERENCES

1. Cressie, N.A.C. (1993). *Statistics for Spatial Data*, Wiley, Revised version.
2. Dempster, A.P. and Schatzoff, M. (1965). Expected Significance Level as a sensitivity index for test statistics. *Journal of the American Statistical Association*, **60**:420-436.
3. Fuentes, M. (2001). A high frequency kriging approach for nonstationary environmental processes. *Environmetrics*, **12**:469-483.
4. Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika*, **89**. To appear.
5. Fuentes, M. and Smith, R. (2002). A new class of models for nonstationary processes. Submitted.
6. Rao, C.R. (1973). *Linear Statistical Inference and its Applications*. 2nd Edition, Wiley, New York.
7. Renshaw, E. (2002). Two-dimensional spectral analysis for marked point processes. *Biometrical Journal*, **6**:718-745.
8. Sampson, P. and Guttorp, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of American Statistical Association*, **87**:108-119.
9. Stein, M. (1999). *Interpolation of Spatial Data: some theory for kriging*. Springer-Verlag, New York.

OPTIMIZATION OF AN ESTUARINE MONITORING PROGRAM: SELECTING THE BEST SPATIAL DISTRIBUTION

S. Caeiro[1], L. Nunes[2], P. Goovaerts[3], H. Costa[4], M.C. Cunha[5], M. Painho[6], L. Ribeiro[7]

[1]IMAR, Depart. of Exact and Technological Sciences of the Portuguese Distant Learning University, .R. Escola Politecnica, 147, 1200 Lisbon, Portugal. scaeiro@univ-ab.pt; [2]CVRM, Faculty of Marine and Environmental Sciences, University of Algarve, Portugal, Campus de Gambelas, 8000 Faro, Portugal, lnunes@ualg.pt; [3] Biomedware, Inc. 516 North State Street, Ann Arbor MI 48104, USA, goovaerts@biomedware.com; [4]IMAR, Faculty of Science and Technology of the New University of Lisbon, Quinta da Torre, 2829-516 Caparica, Portugal, mhcosta@fct.unl.pt; [5]Department of Civil Engineering, University of Coimbra, Pinhal de Marrocos, 3030 Coimbra, mccunha@dec.uc.pt; [6]ISEGI/CEGI, Institute for Statistics and Information Management of the New University of Lisbon, Campus de Campolide, 1070 – 312, Lisboa, Portugal, painho@isegi.unl.pt; [7]CVRM, Lisbon Technological Institute, Technical University of Lisbon, R, Rovisco Pais,1096 Lisboa Codex, nlrib@alfa.ist.utl.pt

Abstract: Monitoring estuarine programs are fundamental to evaluate pollution abatement actions, fulfillment of environmental quality standards and compliance with permit conditions. Their sampling designs should provide statistically unbiased estimates of the status and trends with quantitative confidence limits on spatial scale. The aim of this work is to select a subset of monitoring sampling stations based on locations from an extensive sediment campaign (153 sites) in the Sado estuary (Portugal). In each location three sediment parameters were determined with the objective of defining spatially homogenous environmental areas. The new monitoring program is based on fewer and on the most representative monitoring stations inside each homogeneous environmental area for their future contaminant assessment. Simulated annealing was used to iteratively improve on the mean square error of estimation, by removing one station at a time and estimating it by indicator kriging using the remaining stations in the sub-set, within a controlled non-exhaustive looping scheme. Different sub-set cardinalities were tested in order to determine the optimal cost-benefit relationship between the number of stations and monitoring costs. The model results indicate a 60 station design to be optimal, but 17 additional stations were added based on expert criteria of proximity to point sources and characterization of all homogenous areas.

X. Sanchez-Vila et al. (eds.), geoENV IV – Geostatistics for Environmental Applications, 355-366.

Key words: Optimization, monitoring sampling, indicator kriging, estuarine sediments.

1. INTRODUCTION

Estuaries are coastal transitional water bodies with natural resources of high preservation values, providing important habitats for different species of organisms. The uses inside the estuary and around it have impacts on the water and sediment quality that may put at risk the equilibrium of the ecosystem. Environmental management of these ecosystems cannot be conducted effectively without reliable information on changes in the environment and on the causes of those changes. Ecological monitoring programs can represent an important source of that information. However many of the existing programs are not effective. To assure effectiveness, monitoring programs should be well designed, to enable the statistical analysis and interpretation needed to relate cause and effects (Olsen *et al.*, 1999 and Vos *et al.*, 2000).

The reliability of the sampling design depends on such a large degree on the sampling spatial distribution and size that their importance should not be underestimated (Haining, 1990). One or more of the following principles could govern the size of the sample (Cochran 1977; Clark and Hosking 1986; Strobel *et al.*, 2000):i) the required sampling size can be found if we have reasonable estimates of the population variance measured through a preliminary pilot survey; ii) certain statistical tests require a reasonable sample size; although no fixed minimum can be stated, a sample size of at least 30 is usually employed; iii) too large sample implies a waste of resources, and too small diminishes the utility of the results; iv) finance and time may dictate a certain maximum sample size.

In ecosystems like estuaries the spatial variability of key ecological indicators could be a measure to determine the appropriate monitoring sampling design (Strobel *et al.*, 2000).

The kriging interpolation is very useful to minimise the estimation variance for any fixed sampling design. The plot of the maximum value of the minimised estimation variance against sampling interval, or sample size, can be used to select sample size to achieve a required level of precision (Haining, 1990). For operational, economic or political reasons sometimes sampling sites for monitoring must be reduced and resource allocation optimized (Cochran, 1977). Optimal sampling scheme can then be designed by deleting sites from a current network so as to minimize the variance of estimation error, which means deleting the site that can be predicted best from the remaining sites (Cressie, 1993). Clever search algorithms like

simulated annealing can then help designing the best sampling scheme. Difficulties usually arise in finding an optimal sampling plan and optimal kriging weights. Sampling plans can be important factors when looking for optimal spatial designs. Using the mean-squared prediction error of predictors, the rate of convergence to zero is faster for stratified random sampling than random or systematic random sampling designs (Cressie, 1993).

The sampling optimality criteria should not only be statistical but also cost related or economical (Cochran, 1977, Cressie, 1993, Vos *et al*, 2000). Sampling and parameters measurement costs are very important limitations and should be taken into account in the optimization procedure.

The aim of this work is to select, due to budget constrains, a subset of monitoring sampling stations from an extensive stratified random campaign of estuarine sediments. This subset will be used to assess Sado Estuary sediment contamination in management areas previously delineated. Spatial simulated annealing was used to optimize the sample locations. These data will be further integrated in an environmental management system for Sado Estuary.

2. CASE STUDY

The Sado Estuary, located in the West Coast of Portugal, is the second largest in Portugal with an area of approximately 24,000 ha. The estuary comprises the Northern and the Southern Channels, partially separated by intertidal sandbanks. Most of the water exchange is made through the southern Channel. The estuary is linked to the ocean by a narrow and deep channel that makes a major contribution to the general pattern of the estuarine circulation (Neves, 1986). Most of the estuary is classified as a Nature Reserve. There are many industries mainly on the northern margin of the estuary. Furthermore the harbour associated activities and the city of Setúbal along with the mines on the Sado watershed also releases contaminants into the estuary. In other areas around the estuary, intensive farming, mostly rice fields, is the main land use together with traditional saltpans and increasingly intensive fish farms. Most of these activities have negative impacts on water, sediment and biotic communities namely because they discharge to the estuary contaminants like heavy metals, or organic compounds (Caeiro *et al.*, 2002b).

3. METHODS

3.1 Sediment Homogenous Areas Delineation

In a first extensive campaign 153 sediment locations were sampled for analysis of properties of general characterisation: fine fraction (FF), organic matter (OM), and redox potential (Eh). These key ecological parameters explain main variations in the type and behaviour of benthic organisms as well as contaminant mobility/accumulation (Rodrigues and Quintino, 1993). One method of determining sample size for multiple parameters assessment, is to specify margin error for the items that are regarded as most vital to the survey (Cochran, 1977). A systematic unaligned sampling design with a grid size equal to 0.365 km^2 was used based on prior information on the spatial variation of sediment granulometry (Figure 1) (Caeiro *et al.*, 2002a).

This extensive campaign was intended to help defining homogeneous areas (future management areas) for Sado Estuary within which contamination would be monitored using smaller sample sets.

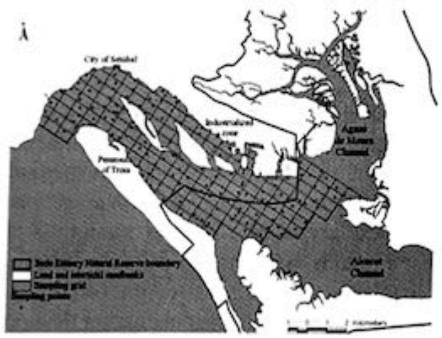

Figure 1. Sado Estuary sediment sampling design (Adapted from Caeiro *et al.*, 2002a).

These homogenous areas were delineated in 5 steps based on grouping individual sampling sites that have similar physicochemical properties while being geographically close (Caeiro *et al.*, submitted): 1)Principal component (PC) extraction of the 3 sediment properties variability (FF, OM and Eh); 2)Variogram fitting of a spherical model to 1st PC factor scores; 3)Dissimilarity matrix determination; 4)Cluster analysis using the complete linkage rule on the dissimilarity. matrix to estimate the probability of occurrence of four selected clusters at sampled stations; 5)Indicator kriging to interpolate these probabilities at unsampled stations; 6)Maximum likelihood classification of these unsampled stations.

The dissimilarity between any two sampling sites i and j (step 3) was computed following Oliver and Webster (1989) equation with spherical model adjustment (Goovaerts, 1997) to take into account the form of spatial variation. Step 5, started with an indicator coding of classification results (x_α) at each sampled station x_α:

$$i(x_\alpha;z_l)=\begin{cases} 1 \ \text{if} \ z(x_\alpha)=z_l \\ 0 \ \text{otherwise}\end{cases} \quad l=1,\ldots,L \tag{1}$$

where L is the number of clusters (four selected). For each cluster z_l, experimental indicator variograms are then computed and modelled:

$$\gamma(\mathrm{h};z_l)=\frac{1}{2N(\mathrm{h})}\times\sum_{\alpha=1}^{N(\mathrm{h})}\left[i(x_\alpha;z_l)-i(x_\alpha+\mathrm{h};z_l)\right]^2 \tag{2}$$

The probability of occurrence of the l-th cluster at the unsampled station x is estimated as a linear combination of indicator data:

$$\hat{p}(x;z_l\,|\,B)=\sum_{\alpha=1}^{n_c}\lambda(x_\alpha;z_l)\times i(x_\alpha;z_l) \tag{3}$$

where B is the set of n_c surrounding data $\{z(u_\alpha), \alpha=1,\ldots, n_c\}$. The weights $\lambda(x_\alpha;z_l)$ are solutions of an indicator kriging system and account for data configuration and spatial continuity of clusters as modelled by indicator variograms. In theory, indicator cokriging estimator is better than the indicator kriging estimator because it accounts for additional information available across categories. However, indicator cokriging improves little over indicator kriging according to Goovaerts (1994).

3.2 Optimization model

The stations that produce the lowest estimation error variance, estimated using cross-validation technique (Deutsch and Journel, 1998), result in a spatial distribution with the highest accuracy. The objective function considers a set, S, of all the original stations, with cardinality Ω, and take a subset, S', with cardinality ω, such that $\omega < \Omega$.

Minimize

$$s^2_{fp} = \frac{1}{\omega}\sum_{\alpha=1}^{\omega}\left[i(x_\alpha;z_l)-i^*(x_\alpha;z_l)\right]^2, \omega \in S', S' \subset S \tag{4}$$

Subject to:

$$\Psi_{S'}(A;z_l) \approx \Psi_S(A;z_l) \tag{5}$$

s^2_{fp} is the mean squared error of estimation and equal to the variance of the estimation error if zero mean estimation errors are considered (i.e. no bias). $i^*(x_\alpha,z_l)$ is the indicator kriging estimated value, $\psi_S(A,z_l)$ and $\psi_{S'}(A,z_l)$ are the marginal probabilities of finding stations with values in $]z_{l-1}, z_l]$ in the original data set and in the candidate solution, respectively.

The new design S' must reflect the sediment physical and chemical variability detected with the prior sampling campaign. Therefore we imposed the constraint that the proportions of monitoring stations in each of the identified homogeneous areas are similar to the proportions in the original sampling campaign (Table 1). Van Groenigen *et al.* (2000) also successfully used sampling constraints in spatial annealing to optimise sampling scheme. The condition is not equality because, for practical computation, floating-point variables equality is machine dependent and varies with the precision. Instead, $\Psi_{S'}(A,z_l)$ may be bounded, and the constraint becomes:

$$\Psi_S(A;z_l)(1-\delta) \le \Psi_{S'}(A;z_l) \le \Psi_S(A;z_l)(1+\delta) \tag{6}$$

A conditioning on the objective function with $\delta = 0.3$ was imposed. This condition is necessary to correct the bias introduced by variogram models fitting errors (when adjusting the theoretical models to the experimental variogram). If no conditioning is used increasing the number of stations will result in higher estimation error variances. This is due to the fact that at very low ω only stations with low estimation error in the optimal solution are included; as ω increases higher estimation error stations are included (Nunes *et al.*, unpublished).

Simulated Annealing (SA) algorithm with the Metropolis iterative improvement procedure (Metropolis *et al.*, 1953) was then used to solve the optimisation model. This procedure generalises by incorporating controlled uphill steps (to worse solutions). The procedure states the following: consider one small random change in the system at a certain temperature (the control parameters t is usually termed temperature); the change in the objective function is ΔOF; if $\Delta OF \le 0$, then the change in the system is accepted and the new configuration is used as the starting point in the next

step; if $\Delta OF > 0$ then the probability that the change is accepted is determined by $P(\Delta OF) = exp(-\Delta OF/t)$; a random number uniformly distributed in the interval (0,1) is taken and compared with the former probability; if this number is lower than $P(\Delta OF)$ then the change is accepted. The SA algorithm runs in the following way: i) the system is *melted* at a high temperature (initial temperature, t_l); ii) the temperature is decreased gradually until the system *freezes* (no further OF change occurs); iii) at each iteration the Metropolis procedure is applied; iv) if any of the stopping criteria is reached the algorithm is stopped and the best solution found is presented. The generic SA algorithm for a minimisation, considering a neighbourhood structure N, a current solution X, a best solution found so far X_{best}, a solution space χ, a α temperature decrease control parameter and an objective function OF has the following pseudo-code.

```
Select an initial solution X_best;
Select an initial temperature t_l>0;
Select a temperature reduction factor;
Repeat
  Repeat
    Randomly select X∈ N(X_best);
     ΔOF = OF(X) – OF(X_best);
     IF ΔOF<0 then
     X_best = X
       else
    generate random z uniformly in (0,1);
           if z < exp(-ΔOF/t) then X_best = X;
  Until iterations = max_iterations
  Set t = αt;
Until stopping condition = true;
X_best is the optimal solution found.
```

In order to speed-up the process several improvements have been proposed, namely by limiting the number of iterations at each temperature, i.e., defining the number *max_iterations*. The dimension of the Markov chain has been proposed to be a function of the dimension of the problem (Kirkpatrick *et al.*, 1983): temperature is maintained until 100Ω solutions (iterations), or 10Ω successful solutions have been tested, whichever comes first. Ω stands for the number of variables (stations) in a problem.

A specific computer code in FORTRAN that incorporates both the estimation error variance and the SA algorithm was developed by (Nunes *et al.*, unpublished) to optimise location problems and adapted to this specific problem. Runs were made on PC Intel 2000 MHz machines.

Fourteen different monitoring network dimensions (cardinality of S': ω) were tested, {25,30,35,40,45,50,60,70,80,90,100,110,120,130} according to the following scheme: i) impose a number of monitoring stations (ω) to be included in the new design; ii) find the optimal allocation solution with SA; iii) increase ω and return to i). SA solutions are considered optimal when more than 70% out of 20 consecutive runs with the same objective function conditions (ω, δ) and SA parameters have the lowest and equal s_{fp}^2 value.

A complementary analysis comparing the loss in accuracy versus reduction in exploration costs as stations are removed was also performed. For that purpose a cost per sampling was computed based on the previous sampling campaign and laboratory analysis costs (official costs of the laboratory where the analysis are going to be made - ControLab, lda.): i) linear distance between n sampling point: n/study area (56 km^2); ii) boat velocity: 12,8 km^2; iii) hour of work per day:7 h/day; iv) time for sampling: 20 min; v) Boat cost per day: 250 Euros; vi) Cost per total contaminant analysis: 500 Euros (discount: 25 % from 20 to 50 stations, 30 % from 55 a 100 stations and 40 % from 105 to 135 stations).

4. RESULTS AND DISCUSSION

Table 1 lists four different physical and chemical homogeneous areas (clusters) based on the sampling campaign data and results from hierarchical classification (step 4), and their frequencies in the study area.

For each cluster, the indicator variogram was computed along four directions and a geometric anisotropic model was fitted (Figure 2).

Table 1. Physical and chemical parameters of each cluster and their frequency.

Sediment Parameter	Clusters (s) High organic load (z_1)	Medium high organic load (z_2)	Medium organic load (z_3)	Low organic load (z_4)
OM (%)	8.6 ± 2.4	4.2 ± 1.4	1.9 ± 0.7	0.9 ± 0.3
FF (%)	60.4 ± 27	21.7 ± 11.8	9.1 ± 7.8	1.5 ± 1.3
Eh (mV)	-278.9 ± 68.6	-178.8 ± 72.6	-137.4 ± 50.9	74.4 ± 49
Freq. (%)	11.76	37.91	23.53	26.80

Figure 3 shows the spatial accuracy plotted versus the monitoring network dimension. Beyond 60, each new added station had little effect on the monitoring spatial accuracy (s_{fp}^2). Sixty is therefore considered as the optimal ω value. The resulting network was overlaid on the sediment homogenous areas within the estuary coast line (Caeiro *et al.*, 2002a) using Arcview/arcinfo 3.2 GIS software (Figure 4a). In cluster one and two (z_1 and z_2) the estimation errors are higher, therefore leading the optimisation algorithm to select preferentially the two remaining clusters with lower estimation errors. These clusters are therefore more densely sampled than in the original data set, as a way to compensate for the bias introduced. Also when high or low values of a cluster are grouped in small areas scattered in

the study area, their relative frequencies are low or data values is too random, the variogram fitting becomes difficult and prone to error. The result is the fitting of theoretical variograms that only roughly approximate the real variability and large estimation errors. This does not hinder the geostatistical method, but justifies the need to impose reproduction of the original proportions (Nunes *et al.*, unpublished).

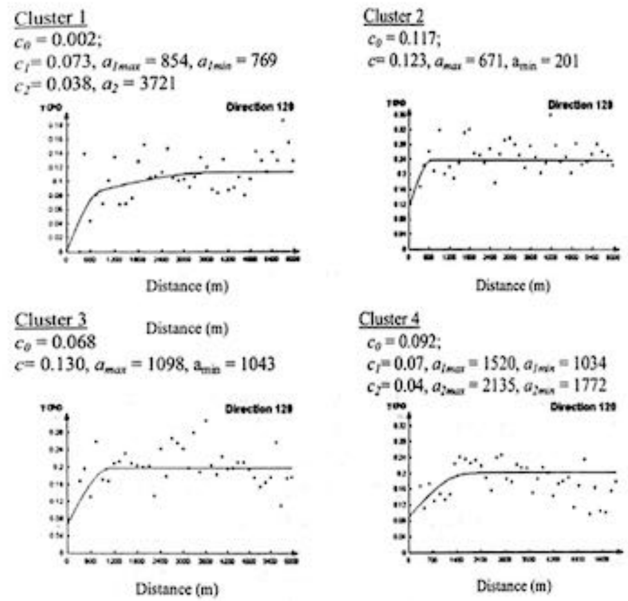

Figure 2. Cluster experimental directional variograms and spherical model fitted for 120°, the major direction of anisotropy. Other directions (not shown) included 30°, 75° and 165°.

Figure 4a) indicates that not all the homogenous areas are sampled in the optimal scheme solution, in particular areas belonging to clusters with high organic load (1 and 2), for the reasons explained earlier. Most of these cluster 1 and 2 areas are near contaminant point sources, mainly in the North Channel. Thus 17 stations were added to the optimal ω value according to expert knowledge aiming to characterize the impact of those point sources and homogenous areas not included in the optimised network (Fig. 4b).

The number of stations to evaluate contamination in the study area (77 stations/56 km^2, corresponding to 1.38 stations/km^2) is within the average of sediment sample size of Environment Monitoring Assessment Program (EMAP) of United States Environmental Protection Agency (USEPA) for small estuaries. The sample sizes for the different estuaries of EMAP vary from 0.11 to 4.16 stations/km^2 (Strobel *et al.*, 2000). Such a wide interval might be related to the spatial variability of sediment parameters in each coastal zone, which is caused by differences related to geomorphological, biological and human pressures.

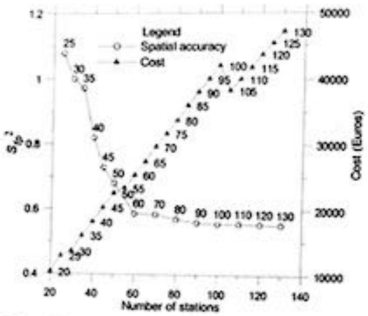

Figure 3. Estimation error variance and cost versus number of monitoring stations.

The exploration costs analysis (Figure 3) showed that costs are always increasing and only for large number of stations (from 110 to 115) does the cost decrease. Indeed the cost of contamination concentration analyses has a high weight in the total cost and only for 105 laboratory analysis does the laboratory discount significantly affect the total cost.

Although seventy-seven stations still represent a high cost (about 60 % of stations total number cost), this budget figure is considered necessary at the present time for a contamination assessment. For any future long-term monitoring program to assess estuary ecological condition, a reduced number of sampling sites should be chosen. Thirty sampling stations should represent a good number for a monitoring program since: i) each of the 19 management areas could be sampled at least at one location or two in case of larger areas, ii) it is a statistical minimum required; iii) the cost is not too high (and similar to 25 stations – see Figure 3). Nevertheless, 30 stations will represent a 40 % loss in spatial accuracy (see Figure 3).

In the future developments for a monitoring program of the environmental management system of the estuary, the model should take into account two strata in the study area. One in the North Channel near pollution sources and the other in the South where the hydrodynamics is highest and the pollution sources are non-point. Vos *et al.*, (2000) discuss that the identification of relevant subsystems or strata for monitoring purpose, is very important to maximise diagnostic of ecological changes. In these strata changes in the anthropogenic inputs or "controlled variables" are expected. Also, once contaminants have been measured at the 77 sampling points a new optimisation criterion could be developed to sample preferentially areas with high priority (e.g. high concentrations). Van

Groenigen *et al.* (2000) used a spatial weight function in spatial simulated annealing that allows distinguishing between areas with different contamination priorities. This could be achieved through Weighted Mean of Shortest Distance; i.e. the fitness is extended with a location-dependent weighing function, or/and using probability maps of contamination and indicator kriging. In particular in our case the weight function should take into account small areas and distance to contaminant sources.

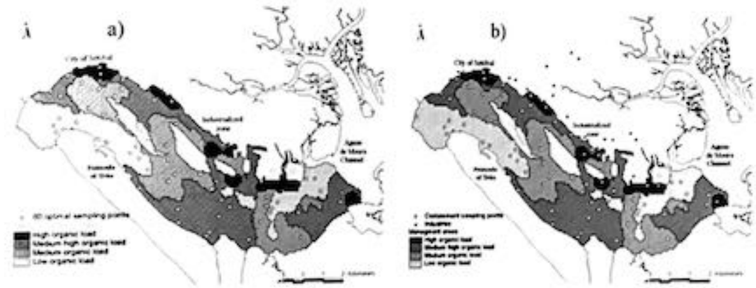

Figure 4. Monitoring networks a) for ω value = 60 stations; b) with 60 optimal stations and additional expertise criteria (17) (Location of industries from Araujo *et al.* (2002).

5. CONCLUSIONS

Monitoring programs should be planned in order to provide quantitative and scientific assessments of pollutants' complex effects on these systems. Optimal sampling designs for ecological condition assessment should take into account not only statistical criteria but also historical knowledge about the study area. In particular estuaries have always areas with different priorities (e.g. human pressures or more sensitive areas). From an extensive campaign including 153 sampling points, a sampling design with 77 stations was selected for sediment contaminant assessment in Sado estuary. This selection was based on minimization of indicator kriging mean square error estimation and expertise knowledge. For a future long–term monitoring program of the estuary condition assessment a reduced subset of 30 stations should be chosen based on definition of contaminant priority areas.

ACKNOWLEDGEMENTS

The two first authors had a PRODEP scholarship. The research was partially financed by the Portuguese Science and Technology Foundation (Research Project BSE/35137/99-00).

REFERENCES

1. Araujo, R., Vasconcelos, L. and Painho, M. SADIND - Sistema de Visualização Interpretativa para a Gestão Ambiental, *Biologia* (In Press), 2002.
2. Caeiro, S., Goovaerts, P., Painho, M., Costa, M. H. and Sousa, S. Optimal spatial sampling design for mapping estuarine sediment management areas. Ruiz, M., Gould, M., and Ramon, J. (Ed.) 5th AGILE Conference on Geographic Information Science; Palma, Spain. Universitat de les Illes Balears, 389 – 396, 2002a.
3. Caeiro, S., Painho M., Costa, M. H. and Ramos, T. B. Sado Estuary Ecosystem: a management methodology. Fernando Pessoa University. Duarte, P. (Ed.) Proceedings of International Conference on Sustainable Management of Coastal Ecosystems; Porto, Portugal. Fernando Pessoa University, 2002b.
4. Clark, W. and Hosking, P. *Statistical Methods for Geographers*. John Wiley & Sons, 1986.
5. Cochran, W. *Sampling Techniques*. 3rd edition ed. New York: John Wiley & Sons, 1977.
6. Cressie, N. *Statistics for Spatial Data*. Revised Edition. John Wiley & Sons, 1993.
7. Deutsch, C. and Journel, A. G. GSLIB. *Geostatistical Software Library and Users's Guide*. 2nd edition. Oxford University Press, 1998.
8. Goovaerts, P. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Mathematical Geology*, 1994; 26(3):389-411.
9. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*: Oxford University Press, 1997.
10. Haining, R. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, 1990.
11. Kirkpatrick, S., Gellat, Jr. C. D. and Vecchi, M. P. Optimization by simulated annealing: Science, 1983; 220:671-680.
12. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. Equation of state calculations by fast computing machines: *The Journal of Chemical Physics*, 1953; 21: 1087-1092.
13. Neves, R. J. J. Hidrodynamical Modelling as a Toll in Waste Disposal Selection. A Case Study on Sado Estuary. Kullenberg (ed.). *The Role of the Oceans as a Waste Disposal Option*, 1986, pp. 563-576.
14. Oliver, M. A. and Webster, R. A Geostatistical Basis for Spatial Weighting in Multivariate Classification. *Mathematical Geology*, 1989; 21(1):15-35.
15. Olsen, A. R., Sedransk, J., Edwards, D., Gotway, C. A., Liggett, W., Rathbun, S., Reckhow, K., Young, L. Statistical issues for monitoring ecological and natural resources in the United States. *Environmental Monitoring and Assessment*, 1999; 54:1 – 45.
16. Rodrigues, A. and Quintino, V. Horizontal Biosedimentary Gradients Across the Sado Estuary, W. Portugal. *Netherlands Journal of Aquatic Ecology*, 1993; 27(2-4):449-464.
17. Strobel, C., Paul, J., Hughes, M., Buffum, H., Brown, B. and Summers, K. Using Information on Spatial Variability of Small Estuaries in Designing Large-scale Estuarine Monitoring Programs. *Environmental Monitoring and Assessment*, 2000; 63:223-236.
18. Van Groenigen, J. W., Pieters, G. and Stein, A. Optimizing spatial sampling for multivariate contamination in urban areas. *Environmetrics*, 2000; 11:227-244.
19. Vos, P., Meelis, E. and Keurs, W. J. A framework for the design of ecological monitoring programs as a tool for environmental and nature management. *Environmental Monitoring and Assessment*, 2000; 61:317-344.

GEOSTATISTICAL ANALYSIS OF THREE DIMENSIONAL CURRENT PATTERNS IN COASTAL OCEANOGRAPHY: APPLICATION TO THE GULF OF LIONS (NW MEDITERRANEAN SEA)

P. Monestiez[1], A. Petrenko[2], Y. Leredde[2] and B. Ongari[2]

[1]*Unité de Biométrie, INRA, Domaine St Paul, Site Agroparc, 84914 Avignon Cedex 9, France.*
[2]*Centre d'Océanologie de Marseille, LOB, Campus de Luminy, Case 901, 13288 Marseille Cedex 9, France*

Abstract: Two geostatistical methods are used to map hydrodynamic patterns in the Gulf of Lions (Mediterranean Sea). The aims are both methodological – mapping vectorial data raises some difficulties – and applied – sampling schemes from boat cruise are not convenient to get maps or to compare with model output. From a large data set that was obtained from a shipboard ADCP (Acoustic Doppler Current Profiler), stationary isotropic geostatistical models were fitted for several horizontal layers. Vectors of current are characterized by two components or by intensity and direction. A linear model of coregionalization was used on vector components and compared to a second approach that considers vectors as elements of the complex plane $\mathbb{C}$. Then interpolated maps were computed by ordinary cokriging and by ordinary kriging in the complex plane for two different depths. Although some difficulties remain unsolved due to the effect of time in the sampling scheme or to some constrains (physical equations and limit conditions) that currents must satisfy, the first results are already satisfactory and allow a better understanding of spatial patterns than the simple plots of original data. The same data set were also used in parallel for hydrological modelling using a physical circulation model. Then the complex kriging approach was used to address the spatial analysis of the residuals, i.e. difference between predicted and observed current vectors. Residuals were highly structured in space.

X. Sanchez-Vila et al. (eds.), geoENV IV – Geostatistics for Environmental Applications, 367-378.