Available online at www.sciencedirect.com
ScienceDirect
Computers, Environment and Urban Systems
30 (2006) 880–904
Computers, Environment and Urban Systems
www.elsevier.com/locate/compenvurbsys

Spatial analysis in ecological risk assessment: Pollutant bioaccumulation in clams *Tapes philipinarum* in the Venetian lagoon (Italy)

Stefania Bertazzon [a,*], Christian Micheletti [b], Andrea Critto [b], Antonio Marcomini [b]

[a] *Department of Geography, University of Calgary, 2500 University Dr. NW, Calgary, AB, Canada T2N 1N4*
[b] *Department of Environmental Science, University Ca' Foscari of Venice, Dorsoduro 2137, 30123 Venice, Italy*

Received 24 July 2005; received in revised form 23 September 2005

Abstract

Exposure characterization is a central step in Ecological Risk Assessment (ERA). Exposure level is a function of the spatial factors linking contaminants and receptors, yet exposure estimation models are traditionally non-spatial. Non-spatial models are prone to the adverse effects of spatial dependence: inflated variance and biased inferential procedures, which can result in unreliable and potentially misleading models. Such negative effects can be amended by spatial regression modelling: we propose an integration of geostatistics and multivariate spatial regression to compute efficient spatial regression parameters and to characterize exposure at under-sampled locations. The method is applied to estimate bioaccumulation models of organic and inorganic micropollutants in the tissues of the clam *Tapes philipinarum*. The models link bioaccumulation of micropollutants in clam tissue to a set of environmental variables sampled in the lagoon sediment. The Venetian lagoon case study exemplifies the problem of multiple variables sampled at different locations or spatial units: we propose and test an effective solution to this common and serious problem in environmental as well as socio-economic multivariate analysis.

Keywords: Ecological Risk Assessment (ERA); Spatial analysis; Multivariate spatial regression; Geostatistics; Edible organisms; Contaminated sediment; Sampling

* Corresponding author. Tel.: +1 403 220 7355; fax: +1 403 282 6561.
E-mail addresses: bertazzs@ucalgary.ca (S. Bertazzon), mikhr@unive.it (C. Micheletti), critto@unive.it (A. Critto), marcom@unive.it (A. Marcomini).

0198-9715/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2005.09.003

1. Introduction

The estimation of exposure levels both in abiotic (sediment) and biotic (tissue) media is a central step in Ecological Risk Assessment (ERA), to estimate the ecotoxicity risks posed by chemicals (US-EPA, 1992). Although, in general, regulatory guidelines emphasize that exposure conditions are a function of spatial factors linking contaminants and receptors (e.g., distribution of sources, distribution of contaminants, and size and location of habitats) (US-EPA, 1998), exposure estimation methods frequently assume a continuous and static exposure to contaminants, described by statistics such as mean or Upper Confidence Limit of the mean (UCL) 95% (Linkov, Burmistrov, Cura, & Bridges, 2002). Even though such estimation methods are conservative, they ignore site-specific conditions that can and should be accounted for in risk assessment. For organisms such as clams, living in association with contaminated sediments, bioaccumulation of organic and inorganic chemicals is likely to be the primary source of risk (Bertazzon, Carlon, Critto, Marcomini, & Zanetto, 2000). The pollutant bioavailability is affected by local sediment characteristics, e.g., organic matter fraction (Griffin, 1991; Mahony et al., 1996).

Non-spatial models are prone to the adverse effects of spatial dependence in the data, primarily an inflated model variance. Not only does a large variance hamper the reliability of the estimates, it also inflates inferential procedures, with the result that irrelevant variables may appear to be statistically significant. Estimates derived from such models are consequently not only unreliable, but potentially misleading. The use of explicitly spatial analytical techniques in the environmental sciences and specifically in ERA is still fairly uncommon (US-EPA, 1998), nevertheless, these techniques have the ability to overcome these major flaws of standard, non-spatial procedures.

Our proposed method for exposure characterization hinges on spatial regression, which provides efficient parameter estimates (i.e., minimal error variance) by taking into account the spatial dependence in the data. The development of our model is based on a preliminary geostatistical analysis (i.e., Kriging), which serves two purposes: to aid in producing an accurate spatial interpolation of the sample points, and to assist in estimating the extent (range) of the spatial dependence in the data. Our exposure analysis is conducted within a GIS (Geographic Information Science) framework, which allows us to integrate different spatial analytical techniques and to effectively visualize our results, enhancing the potential of ERA as a planning and management tool. The goal of this study is the definition and calibration of spatial bioaccumulation models as a step in the development of an ERA framework that explicitly considers the role of spatial factors in the estimation of exposure. Our multivariate bioaccumulation regressions (one for each organic and inorganic micropollutant) model the relationship between bioaccumulation in clam tissues and the environmental characteristics of the sediment. We demonstrate how the coefficients estimated by the regression models can be used to characterize exposure at locations where only sediment variables are known.

We chose a complex case study, characterized by local anomalies and inconstant spatial variability, and by scarcity and clustering of the sampled data: such characteristics, which tend to hamper the validity of statistical models, are rather the norm than the exception in applied ERA. The case study allows us to demonstrate the value of the proposed method: the integration of spatial analytical techniques, to obtain efficient estimation of variables in sampled and under-sampled areas.

Section 1.1 describes the study region and data. Section 2 presents the methods applied: exposure characterization (Section 2.1), and spatial analysis, i.e., geostatistics and multivariate spatial regression (Section 2.2). Section 3 illustrates the application of spatial analytical tools to the case study, discussing the choice of techniques and the underlying hypotheses. Results of geostatistics, spatial regression, and a coefficient-based estimation of exposure at under-sampled locations are also presented in Section 3. A discussion of the methods, case study, and results is presented in Section 4.

1.1. Study area, background, and data

The Venetian lagoon, ca. 550 km^2, hosts a great variety of natural habitats; both the historical center of Venice and the surrounding lagoon (Fig. 1) are UNESCO world heritage sites. Due to its strategic importance as a natural defence and a privileged harbour for the Republic of Venice, a conscious effort to preserve the lagoon began in the Middle Ages: man-made transformations that took place over several centuries have enhanced and at the same time endangered the diversity of the natural environment. Early management practices were mainly of hydraulic nature, aimed at preserving the lagoon from becoming either eroded or filled with river-borne material; these activities culminated in major river diversions in the 16th century and are still on-going. In the early 20th century a chemical industrial plant, Porto Marghera (the second largest of its kind in Europe) was established on the edge of the lagoon; as a consequence, large amounts of organic and inorganic pollutants have been discharged in the lagoon, resulting in sediment contamination.

Through the bioaccumulation process, chemical pollution represents a stressor for the ecosystem, particularly for organisms living in contact with sediments. Moreover, the bioaccumulated chemicals in organisms that live in the sediment are likely to be biomagnified through the food chain, reaching higher concentrations in species at the top level. This raises concerns with respect to human health, since some of those organisms (e.g., clams, as well as other shellfish) are subject to human consumption. Further concerns arise when economic factors are considered, because most edible species are valuable economic resources: the Venice lagoon is, presently, the main producer of clams (*Tapes philipinarum*) in Europe. In this context, several environmental management options, such as the regulation of contaminant concentration in industrial and municipal discharge and the definition of lagoon zones for fishing and aquaculture, are still open.

For the definition and calibration of bioaccumulation models of pollutants in clam tissues, clam data (dependent variable), as well as sediment data (independent variables) were collected. Clam data were obtained from two sources. Dataset A (23 sample points) was collected as part of a project managed by *Consorzio Venezia Nuova* and funded by the Venetian Water Authority (*Magistrato alle acque*); for this project, "Mapping the pollution of the lagoon bed" (1999), data were collected on the concentration of metals (mercury, arsenic, lead, cadmium, copper, zinc, nickel, chromium), as well as organic micro-pollutants (polychlorodibenzo dioxins and furans -PCDD/Fs-, polychloro biphenyls -Total PCBs-, and the sum of dioxin-like polychloro biphenyls congeners -PCBdl-) in the superficial sediments of the lagoon (15 cm depth) and in the tissue of clam (*T. philipinarum*). Dataset B (30 sample points) consists of data on metal concentrations in clam tissue, collected by the Health Authority of the Veneto Region in 1997. After verifying their homogeneity by means of exploratory analyses, datasets A and B were merged,

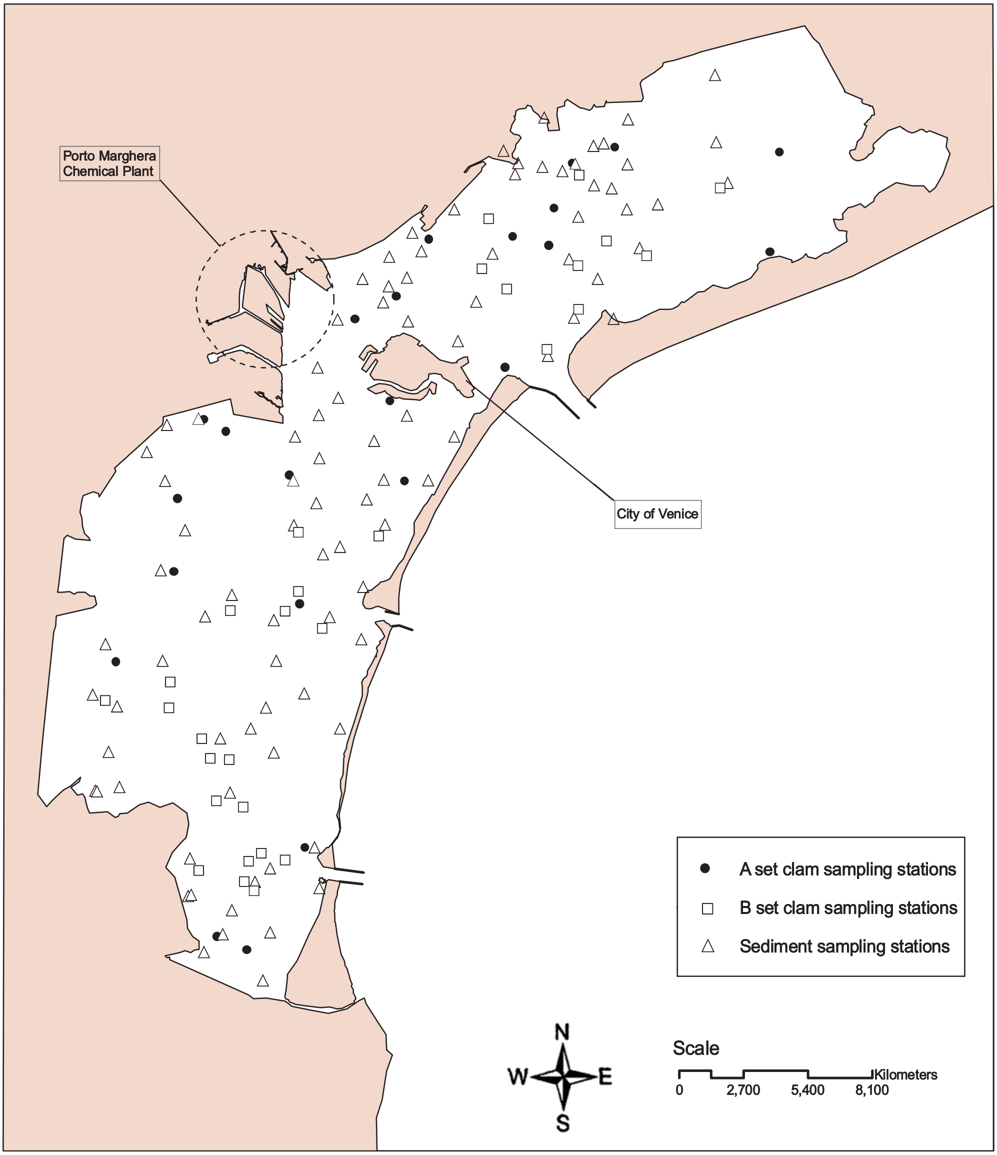

Fig. 1. Overview of the lagoon; clam and sediment sampling stations.

resulting in 53 samples for metals, and 23 samples for organic micropollutants. It should be noted that most of the sampling stations are located in the north and central-south lagoon, where clam farming is currently practiced or planned (Fig. 1). For sediment data, 95 sampling stations are used, which provide information on both metals and organic micropollutants. These samples were also selected by the Venetian Water Authority due to the high quality of the data: the sample plan is consistent and certified by Quality Control and Quality Assurance procedures for sampling and analysis (sample points do not

Table 1
Descriptive statistics of pollutant concentration in clam tissues

Pollutants (mg/kg, w.w.)	Mean	Standard deviation	Minimum	Maximum	Kurtosis	Sample size
Metals						
Arsenic, As	3.8	1.3	1.5	7.4	0.5	53 (A + B sets)
Cadmium, Cd	0.12	0.084	0.008	0.33	−0.45	
Chromium, Cr	0.31	0.07	0.15	0.47	−0.27	
Mercury, Hg	0.05	0.02	0.02	0.11	0.45	
Nickel, Ni	0.79	0.36	0.17	1.87	0.97	
Lead, Pb	0.28	0.18	0.05	0.97	3.8	
Copper, Cu	1.58	0.4	0.89	3	2.3	
Zinc, Zn	15.3	4.26	7.7	29.5	1.6	
Organic compounds						
Sum of dioxin like PCBs congeners	1017	720	51	2555	−0.83	23 (A set)
Total PCBs	5.33	3.5	0.4	13	−0.99	
PCDD/Fs	7.8	12.4	0.53	59	14	

include industrial canals, which would represent an artificial environment whose sediment and contamination characteristics are not comparable with the natural conditions of the lagoon bed). The selected sampling stations are representative of the entire environment, including habitats characterized by different end-uses and environmental characteristics (e.g., contamination, hydrodynamic, etc.). It should be emphasized that small sample sizes and clustered sampling stations are frequent occurrences with this type of data, where collection and analysis are very expensive, and collection campaigns are conducted infrequently and for different purposes.

Selected descriptive statistics of the database used for analysis are presented in Table 1. The software used is Statistica 5 (kurtosis statistic is estimated anchoring the result around zero).

Generally, for purpose of risk assessment, concentration of mixtures of PCBs and PCDD/F congeners in both biotic and abiotic environmental matrices are expressed in Toxicity Equivalents (TEQ) (Im et al., 2004; Van den Berg et al., 1998). However, the TEQ is an expression of toxicity, rather than one of congener bioaccumulability, and for this reason the estimation of regression models based on TEQ can be misleading. Moreover, the analysis of data on individual congener concentrations shows that the relative contribution of individual congeners (in percentage over the sum of congener concentrations) is similar in sediments and clams, suggesting a direct relation between sediments and clams. Once the regression models have been estimated for each congener, the estimated values of congener tissue concentrations can be converted to TEQ to complete the risk assessment procedure.

2. Methods

2.1. Exposure characterization

Our primary objective is the characterization of benthic organism exposure based on site-specific estimation of bioaccumulation. Our method involves the calibration of a multivariate spatial regression model, which will be used to extrapolate estimate values for unsampled locations within the lagoon.

Pollutant bioavailability is a prerequisite for the occurrence of bioaccumulation, and generally only the dissolved fraction of pollutants is fully bioavailable. The main factors that affect bioavailability are: pollutant concentration in the sediment; physical–chemical properties of compounds; sediment characteristics, such as total organic carbon fraction (TOC), fine particle (diameter < 0.063 mm) fraction (ff), and pH; and organism biological characteristics, such as body weight (W) and lipid fraction (L).

For instance, hydrophobic pollutants are strongly associated with organic matter (i.e., total organic carbon) and fine particle fraction (e.g., clay). These properties suggest that an increase in fine fraction and organic matter content in the sediment will result in decreased bioavailability of hydrophobic compounds (US-EPA, 1998). Inorganic pollutant bioavailability is affected by metal speciation, depending mainly on the oxidation state and pH of the sediment (Mountouris, Voutsas, & Tassios, 2002). Biological factors, such as body weight and lipid fraction in organism tissues, are also expected to affect bioaccumulation: weight is an indicator of body growth, which implies dilution of pollutants present in tissues; a greater lipid fraction involves a larger amount of accumulated hydrophobic pollutants (Gobas, 1993).

In kinetic bioaccumulation models, the uptake and elimination rates are estimated by applying allometric equations that link the bioaccumulation of pollutant in an organism to the organism's weight (Mendez, Giudice, Pereira, Inocente, & Medina, 2001), as shown in (1):

$$C_b = aW^b \tag{1}$$

where C_b is the pollutant concentration in the organism's tissues, W is the organism's weight, and a and b are regression coefficients.

This type of model focuses on assessing the organism's uptake and elimination rate, whereas the focus of this paper is on the relationship between bioaccumulation and environmental properties; therefore, we prefer a model formulation that focuses on such relationships. Pollutant concentration in the sediment is expected to be the main factor affecting the organism's exposure: it should be the most significant variable and present a positive sign. Conversely, total organic carbon (TOC), fine fraction (ff) and pH, which tend to reduce pollutant bioavailability, are expected to be negatively correlated with bioaccumulation. We consider multivariate spatial regression models the most appropriate tools, because they provide reliability of estimation parameters (Section 2.2.2), and direct estimation of bioavailability, by integrating exposure with uptake through several pathways (water, sediment, food).

Exposure characterization based on pollutant residuals in organism tissues has been employed with increasing frequency (Munns et al., 1997), also supported by the recent acquisition of residue-effect relationships (Jarvinen, Nordling, & Henry, 1983; Montanes, Van Hattum, & Deneer, 1995; Serrano, Hernandez, Pena, Dosda, & Canales, 1995).

2.2. Spatial analysis

By the term spatial analysis we refer to a set of methods developed by quantitative geographers in order to address the properties of spatial data, particularly those that have an adverse effect on models and estimates. Perhaps the most critical of such properties is spatial dependence, informally described as "near things tend to be more related than distant things" (Tobler, 1979): the presence of spatial dependence in a dataset increases the variance of most statistical techniques, hampering the reliability of the estimates.

Spatial analytical techniques are based on a set of assumptions on the spatial distribution of the phenomenon under consideration, often referred to as a spatial process (Gatrell, 1983). Most spatial analytical techniques, including spatial regression, assume that the spatial process under examination is (1) stationary (Section 2.2.1) and (2) isotropic. In less formal terms, these assumptions require that the effect of distance be (1) constant over the entire study area and (2) direction-invariant. Additionally, multivariate techniques, such as spatial regression, assume these properties for all the spatial processes involved, each process represented by one (dependent or independent) variable. Both assumptions (1) and (2) are rarely met by empirical data, and variables in our study area are no exception.

Indeed, the Venetian lagoon is a complex environment, where a number of factors affect the spatial variability of the sediment and its properties: varying bed depth, discontinuities such as channels and islands, currents, tide, river mouths and winds; a wide range of human activities, including industrial productions, fishing and fish farming, as well as commercial, recreational, and public transit navigation. All these factors, in turn, impact on the distribution and representativity of any data sample. Primarily, it should be noted that data collection and analysis is expensive and collection campaigns are conducted infrequently and for different purposes. Additionally, the apparent irregularities in the spatial distribution of the sample reflect in part local anomalies (e.g., islands or channels), and in part environmental conditions, i.e., samples are clustered in areas that are suitable for clam fishing and/or farming. All these elements have an impact on the properties of the sample, i.e., spatial dependency and stationarity, and ultimately affect the properties of the regression models. Paradoxically, all these features, which may appear as shortcomings of one case study, are common in applied environmental analyses: they do hamper the properties of the regression models and the inferential procedures, but they do, at the same time, point to some of the shortcomings of the currently available analytical tools. It should be reiterated that, even though the assumptions about spatial dependency may still be too simplistic, the inclusion of any kind of assumption on spatial dependency brings the model one step closer to reality than the conventional practice of ignoring spatial dependency entirely.

2.2.1. Geostatistics: Kriging

Geostatistics is a branch of applied statistics that focuses on the spatial context and the spatial relationships among data. It provides tools for the quantification and exploitation of spatial autocorrelation, and algorithms for data interpolation and uncertainty quantification (Isaaks & Srivastava, 1989). A unique aspect of geostatistics is the use of spatial variables that describe phenomena with a geographical distribution (e.g., sediment concentration of pollutants). Even though such phenomena (or spatial processes) exhibit various degrees of spatial continuity, it is not always possible to sample them at every location. Therefore, unknown values must be estimated from data taken at specific locations that can be sampled: geostatistical procedures can be used to estimate values of a surface at the nodes of a regular grid from irregularly spaced data points.

One of the spatial analytical techniques employed in this work is a geostatistical interpolator: Kriging. Unlike in standard inverse distance weighted interpolators, Kriging weights are based not only on the distance among points, but also on their attribute.

The ordinary Kriging estimator is a weighted average of the data values. The weights are obtained as a solution to a system of linear equations; this system is derived by placing two constraints on the estimator: (1) it should be unbiased, (2) the variance of the estimation error should be minimal. The calculation of the Kriging weights is based upon the estimation of a semivariogram model, described by (2):

$$\gamma(h) = \frac{1}{2}\mathrm{Var}[z(s+h) - z(s)] \tag{2}$$

where h is a vector and $h \geqslant 0$ (Haining, 1997). The estimation of a semivariogram model relies on two important assumptions: that the quantity $\gamma(h)$ exists and is finite for all choices of h and s, and that it does not depend on s. The latter property is known as intrinsic stationarity. The Kriging estimator is unbiased and has minimum error variance. The ordinary Kriging estimator is presented in (3):

$$Z(x,y) = \sum_{i=1}^{n} w_i z_i \tag{3}$$

where n is the number of considered measures, z_i are the corresponding attribute values, and w_i are the weights (Isaaks & Srivastava, 1989).

In order to model the variogram, some key features of the empirical variogram are identified that reflect the properties of the spatial process under examination; the process of assessing such properties is known as structural analysis (Davis, 2002). For the present application, the most important model feature is the range, or the distance at which an increase in the separation distance between pairs no longer causes a corresponding increase in the average squared difference between pairs of values and the variogram reaches a plateau (Isaaks & Srivastava, 1989). We interpret the range as an indicator of the extent of the spatial dependence in a process, and use this value in spatial regression models, to define the contiguity threshold that discriminates between spatially dependent and independent observations (Section 2.2.2). Semivariograms and Kriging were computed in GS+ 5.1 (Gamma Design Software, 2001).

2.2.2. Spatial regression analysis

Spatial regression analysis is designed to address the property of spatial dependence, which is commonly present in spatial processes (Anselin, 1998), tending to inflate the variance of standard regression estimates. Even though the concurrent presence of spatial dependence and spatial non-stationarity is common in applied analysis, current regression methods can be grouped in two major categories: spatial autoregressive methods (Anselin, 1988), which specifically address spatial dependence, but typically disregard non-stationarity; and geographically weighted methods (Fotheringham, Brundson, & Charlton, 2002), which address non-stationarity but typically disregard spatial dependence. In our context, estimates of independent variables at un-sampled locations depend strongly on a stationarity assumption. Additionally, it is our goal to estimate exposure values in under-sampled areas using the coefficients of a regression model, which contradicts the emphasis of geographically weighted regression on microscale local relationships. Therefore, we deem spatial autoregressive models most appropriate for this work. Exploratory spatial analyses on our data support this choice: spatial autocorrelation tests confirmed the presence of spatial dependence in most variables, while Monte Carlo simulations indicated that the null hypothesis of stationarity cannot be rejected.

Spatial dependence implies a redundancy of information that increases the variance (uncertainty) associated with the parameter estimates. The presence of spatial dependence in the data affects the model properties by introducing non-null off-diagonal elements in the variance matrix: in the presence of variance larger than the minimum, parameter estimates are inefficient. Large parameter variance inflates classical inferential tests, resulting in a more frequent rejection of the null hypothesis. As a consequence, inefficient parameter estimates are not only unreliable, but potentially misleading. A solution to this problem is provided by spatial regression, through the estimation of an inverse of the variance matrix. The core of the model is an autoregressive procedure, which accounts for the influence on each observation of a number of neighbouring points, weighted by a contiguity matrix. One of the most delicate steps of spatial regression analysis is the specification of an appropriate contiguity, or neighbourhood, matrix (Bertazzon, 2003). The spatial regression equation is described by (4).

$$Y = X\beta + \rho WY + \varepsilon \tag{4}$$

where ρ is the autoregressive parameter and W represents the contiguity matrix. In its simplest form, W is a binary structure, while some more complex specifications include various types of weights that describe distance decay effects (Bailey & Gatrell, 1995).

There are several ways of specifying spatial contiguity (Getis & Aldstadt, 2004). In this study, we propose the use of a semivariogram (Section 2.2.1), as a model of spatial dependence. The semivariogram provides a threshold, or range of spatial dependence, which is used here to define the matrix W as a tool to discriminate between spatially dependent and independent observations. As summarized in (5), observations separated by a distance not greater than the range are defined as spatially dependent, and observations located at a greater distance are defined as independent.

$$\begin{aligned} w_{ij} &= 1 \quad \text{if } d_{ij} \leqslant r \\ w_{ij} &= 0 \quad \text{if } d_{ij} > r \end{aligned} \tag{5}$$

where r is the semivariogram range, and d_{ij} is the distance between observations i and j. While this method is common in geographical analysis, alternatives include spatial models directly based on a variogram-like model for the correlation matrix (Grondona & Cressie, 1991). The regression analysis was performed in Splus 6 and S+SpatialStats 1.5.

3. Application and results

3.1. Kriging interpolation and semivariogram ranges

Geostatistical analysis plays an essential role throughout this study: after using Kriging to interpolate and estimate sediment values, semivariogram ranges are used to define contiguity matrices in spatial regression models; finally, Kriging is used again to interpolate regression model predictions and visualize their results, in order to provide a continuous representation of contaminant distribution from punctual experimental values (Jarvis, Stuart, Baker, & Morgan, 1999; Juang & Lee, 1998; Leonte & Schofield, 1996; Nathanail, Ferguson, & Brown, 1998).

Spatial regression analysis assumes that observations on all variables are taken at the same spatial locations. Datasets that satisfy this assumption are rare in environmental as well as in socio-economic research, and in our case study, most of the clam sampling stations do not coincide with the sediment sampling stations (Fig. 1). Given the larger size of the sediment sample (95 observations vs. 53 observations for metals in clam tissues and 23 for organic pollutants in clam tissues), the distance between clam samples and sediment samples in general is relatively small. Prior to the specification of the regression model, a preliminary Kriging interpolation was performed, to estimate the values of the sediment variables at the clam sample locations. The use of interpolated values for explanatory variables induces an additional element of variability in the regression model, since the estimated values may include interpolation errors. Estimated variables are stochastic by construction, and their use in a regression model violates one of the fundamental assumptions of the model (i.e., independent variables should be non-stochastic). In general, measurement error in explanatory variables may result in non-zero correlation between the variable and the error term, which in turn results in biased and inconsistent estimates. The classical solution to this problem is the use of instrumental or proxy variables, chosen for their high correlation with the explanatory variables and uncorrelated with the error term (Gujarati, 2003). Identifying proxy variables is often a problem, and certainly no such data was available for our work. We chose to use estimated values, and the implications of this choice are discussed in Section 3.2.

The parameters of the semivariogram models for the independent variables are presented in Table 2, while diagrams are plotted in Fig. 2. These models were chosen based on RSS (Residual Sum of Squares) provided by GS+ software. In all cases the model was assumed to be isotropic; we performed block Kriging, assuming a circular search radius, with a range equal to the semivariogram range, and a maximum number of 6 neighbours.

Of the environmental variables, fine fraction (ff) and pH are best fitted by an exponential model, characterized by an almost negligible nugget effect, and similar range values; total organic carbon is fitted by a spherical model, with a more prominent nugget effect, and a larger range. Two of the three organic compounds (PCBdl and PCDD/Fs) are fitted

Table 2
Semivariogram models

Sediment variables	C_0	$C + C_0$	Range	Model
Metals				
Arsenic	4.30	40.84	14,800	Sph.
Cadmium	0.04	0.32	30,180	Exp.
Chromium	0.10	186.10	6150	Exp.
Copper	0.10	115.70	9390	Exp.
Lead	0.10	258.60	5450	Sph.
Mercury	0.05	0.22	20,700	Exp.
Nickel	5.62	36.81	7950	Exp.
Zinc	2020.00	8820.00	15,230	Sph.
Organic compounds				
Sum of dioxin like PCBs	0.42	1.92	5940	Exp.
Total PCBs	24.10	121.53	13,286	Linear
PCDD/Fs	0.27	1.33	8040	Exp.
Sediment characteristics				
Total organic carbon	0.09	0.48	12,020	Sph.
Fine fraction	1.00	693.50	6960	Exp.
pH	0.01	0.07	5610	Exp.

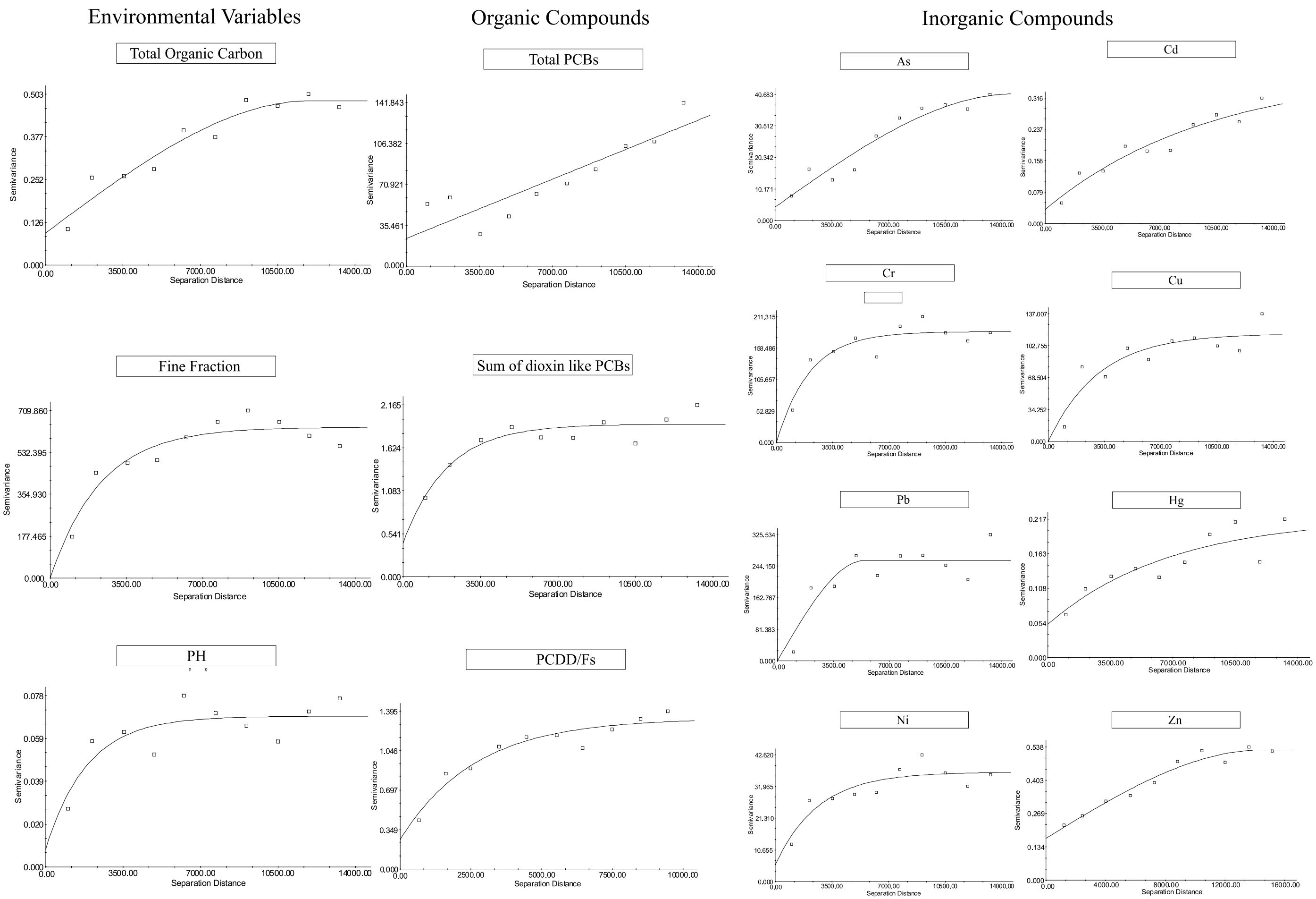

Fig. 2. Variogram diagrams for environmental variables, organic and inorganic compounds.

by exponential models, and Total PCBs by a linear model; all the organic compounds present a non-negligible nugget effect. For six of the eight metals, the best model is exponential, while a spherical model best fits arsenic and lead. Metals tend to present low nugget values, with the exception of zinc and mercury and, to a lesser extent, cadmium and arsenic.

Exponential models suggest a gradual increase of the semivariance as a function of distance, implying that spatial dependence declines at a steady rate with increased distance. This trend is particularly evident in the cadmium and mercury models, which almost approach a linear function; a similar trend is noticeable in the spherical model of zinc. The spherical model of lead presents the most abrupt transition from presence to absence of spatial autocorrelation, while a smoother transition is associated with the larger ranges of zinc and arsenic. The differences in model (e.g., exponential, spherical, linear) and range for the pollutants (i.e. pollutant sediment concentration) and the environmental variables (i.e. pH, TOC, ff), suggest that there are some environmental processes affecting the distribution of pollutants. The main factors could be the type of source (e.g., point or diffuse, natural or artificial), the rate of water exchange in lagoon zones (i.e., confinement degree), and the transport of resuspended sediment (Facca, Sfriso, & Socal, 2002). All these factors simultaneously determine the pollutant distribution in the lagoon environment. For instance, a point source of an organic pollutant (e.g., a river mouth) in combination with a confined area may produce a high sediment concentration in a limited zone of the lagoon, resulting in a short range value and a steep model. On the contrary, a diffuse source of the same pollutant, such as atmospheric deposition (Guerzoni, Rossini, Molinaroli, Rampazzo, & Raccanelli, 2004), may lead to a more homogeneous distribution of pollutant concentration in sediments. Finally, sediment re-suspension and transport distribute over the whole lagoon the organic pollutant adsorbed to organic matter and fine fraction.

Range values represent an indication of the extent of the spatial dependence in each variable, and for this reason they will be used in defining the contiguity matrix for the spatial regression models. Typically, the range of the dependent variable should represent the spatial dependence of the spatial process of interest; in this case study, the range of the variogram of each pollutant in the sediment is used instead of the range of the pollutant in the clam (dependent variable), which can only be estimated on 23 points for organic compounds, and on 53 points for inorganic compounds. However, ranges of clam and sediment variables present a considerable consistency for organic compounds; specifically, for PCDdl the range is 5940 m for the sediment, and 6460 m for the clam.

3.2. Spatial regression

One spatial regression model is calibrated for each micropollutant: the model specification and the expected sign and significance of explanatory variables are discussed in Section 2.1. For inorganic micro-pollutants in clam ($\mathrm{IMP_c}$) the model includes sediment contaminant (SC), total organic carbon (TOC), fine fraction (ff), and pH:

$$\mathrm{IMP_c} = \beta_0 + \beta_1\mathrm{SC} + \beta_2\mathrm{TOC} + \beta_3\mathrm{ff} + \beta_4\mathrm{pH} + \varepsilon \tag{6}$$

For organic micro-pollutants in clams ($\mathrm{OMP_c}$), in addition to the above explanatory variables, the model includes biological characteristics of the organisms, i.e., body weight (W) and lipid fraction (L).

$$\mathrm{OMP_c} = \beta_0 + \beta_1\mathrm{SC} + \beta_2\mathrm{TOC} + \beta_3\mathrm{ff} + \beta_4\mathrm{pH} + \beta_5 W + \beta_6 L + \varepsilon \tag{7}$$

To fully take into consideration the cross-correlations among independent variables, a backwards model selection procedure was applied, and variables were retained or discarded based on the values of their *t*-statistic (Legendre et al., 2002). The estimation method used in S-plus is maximum likelihood: in addition to the standard output we provide a more intuitive goodness-of-fit index, a pseudo-R^2 (Anselin, 1993), computed as the squared correlation between observed and estimated values of the dependent variable. The pseudo-R^2 index is important in Ecological Risk Assessment, because it is used by the United States Environmental Protection Agency (US-EPA) to assess regression models: regression models for bioaccumulation processes are considered acceptable by the US-EPA (2000) when R^2 exceeds 0.20 in value.

The application of spatial regression analysis evidenced a considerable difference in goodness-of-fit between inorganic pollutant models and organic pollutant models. Two of the three organic pollutant models present pseudo-R^2 values above 0.5, and the significance of explanatory variables and the sign of their coefficient are generally consistent and interpretable. The eight inorganic pollutant models present values consistently below 0.25; the set of significant variables varies across models, and the signs of the coefficients are inconsistent. In conclusion, the inorganic pollutant models give inconsistent results, contradictory with the expectations, making their interpretation difficult. Regression models for the bioaccumulation of organic pollutants are presented in Table 3.

The most significant explanatory variables for Total PCBs and the sum of dioxin-like PCB congeners (PCBdl) are sediment contaminant (SC), total organic carbon (TOC), and fine fraction percentage (ff); for PCDD/Fs the significant variables are sediment contaminant and fine fraction. Regression models for organic compounds present relatively high values of pseudo-R^2 (greater than 0.5 for Total PCBs and PCBdl) (Table 3), suggesting that the models are suitable predictors of exposure values. Indeed, in both PCB models the sign and significance of coefficients are consistent with our expectations: total organic carbon (TOC) and fine fraction (ff) are inversely correlated with bioaccumulation of organic pollutants. The most significant explanatory variables are TOC and ff, consistent with the chemical-physical characteristics of hydrophobic organic pollutants, since it is known that the most important factor in controlling the bioavailability of hydrophobic compounds is their adsorption of particulate matter. The role of biological variables (body weight and lipid fraction) is generally negligible, and this can partly be ascribed to the extreme variability of individual characteristics of organisms, also influenced by environmental conditions (e.g., temperature and dissolved oxygen).

An appropriate estimation of the spatial dependence in the data is crucial for the specification of efficient spatial autoregressive models: the contiguity matrix, W, (Eq. (4)) translates the spatial dependence structure into the autoregressive model. For each spatial regression model, the range of the semivariogram of the spatial process of interest (dependent variable) should enter in the contiguity matrix (W) as the maximum distance parameter (w), or the threshold between autocorrelated and uncorrelated spatial units (Eq. (5)). While this criterion provides an indication of the extent of the spatial dependence in one variable (typically the dependent), it does not consider the spatial dependence in any other variable, nor the cross-correlation among variables (Section 3.1). As the structural analysis (Section 3.1) indicates, ranges of pollutants vary substantially, and in some cases they differ considerably from ranges of environmental variables (Table 2).

Table 3
Spatial regression models for organic micropollutants

		Intercept	Pollutant in sediment	Total organic carbon	pH	Fine fraction	Weight	Lipid content	pseudo-R^2	Log likelihood
PCB dioxin-like	Coeff.	2165.97	0.21	−223.47		−7.87			0.54	−174.90
	S.E.	257.7653	0.0327	52.3602		3.9314				
	(t value)		6.41	−4.27		−2.00				
	$p(t)$		0.00	0.00	>0.05	0.06	>0.05	>0.05		
PCB total	Coeff.	9.50	0.27	−1.46					0.52	−56.47
	S.E.	1.0404	0.0523	0.2785						
	(t value)		5.09	−5.26						
	$p(t)$		0.00	0.00	>0.05	>0.05	>0.05	>0.05		
PCDD/F	Coeff.	70.48	0.02			−0.73			0.22	−97.60
	S.E.	6.8796	0.0141			0.0809				
	(t value)		1.72			−8.89				
	$p(t)$		0.10	>0.05	>0.05	0.00	>0.05	>0.05		

To address such potential inconsistencies, we experimented with several *w* values, in order to assess the robustness of the spatial regressions to variations in the specification of the contiguity matrix. The results of this test on the PCBdl model are summarized in Fig. 3, where a few selected indicators are compared. Residual standard error is an indicator of the model variance, the autoregressive parameter, rho, is an indicator of the significance of the autoregressive term, and the logarithm of the likelihood can be read as an indicator of goodness-of-fit of the model. Values in Fig. 3 are standardized and scaled for comparability.

The *w* value of 1500 m is the lowest for which the autoregressive model can be computed in Splus; 12,020 m is the largest range of a variable (total organic carbon) in the model. Additional experiments for *w* values up to 40,000 m show no substantial variation in any of the indicators beyond 15,000 m.

Fig. 3 suggests that the model is robust to variations in the contiguity matrix, as variations in the *w* parameter have only minor effects on the autoregressive term and the error variance, as well as on the model's goodness of fit. This is true only for *w* values beyond a threshold of approximately 6000 m, which can be reasonably interpreted as the range of the pollutant (6460 m in clams, 5940 m in the sediment). The symmetric trends of the autoregressive parameter and the residual standard error also suggest that the autoregressive model effectively reduces the variance in the model estimates. A local optimum (i.e., a minimum of the residual standard error and a maximum of the autoregressive parameter) can be observed at approximately 6000 m, providing strong support for the use of the range of the dependent variable as the distance threshold parameter in the contiguity matrix. A secondary local optimum at approximately 12,000 m can be related to the range of total organic carbon (12,020 m), which is the most significant variable in the regression after the contaminant in sediment. The backwards model selection procedure produces consis-

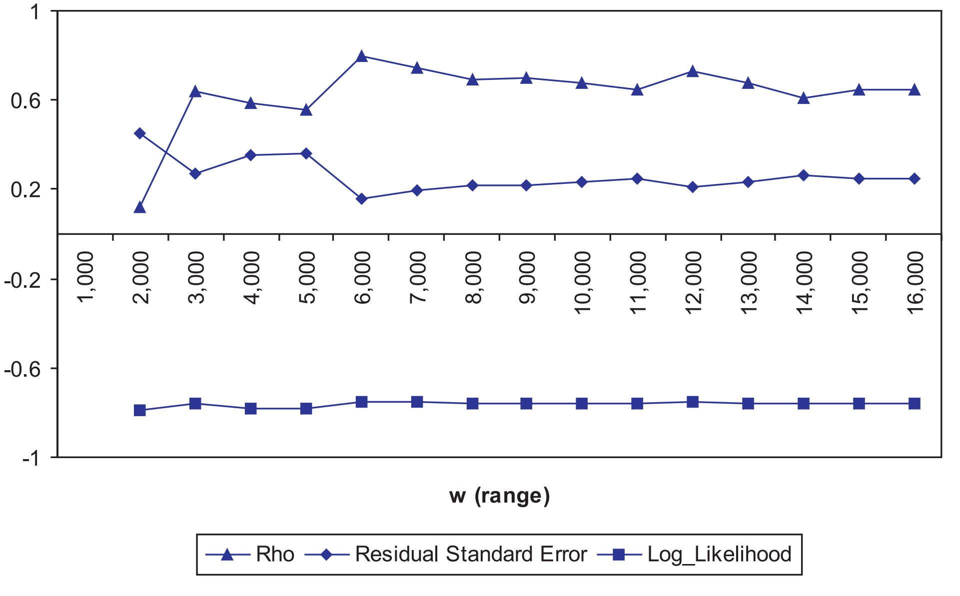

Fig. 3. Model summaries for varying distance threshold in the contiguity matrix.

Table 4
Correlation between model error and independent variable estimation error

	Dioxins		PCB total		PCB dioxin-like	
	Correlation	Spatial regression model residual	Correlation	Spatial regression model residual	Correlation	Spatial regression model residual
Kriging standard deviation	TOC	0.05	TOC	−0.38	TOC	−0.37
	ff	−0.03	ff	−0.43	ff	−0.32
	pH	0.04	pH	−0.50	pH	−0.40
	Dioxins	0.03	PCB total	0.48	PCBdl	−0.42

tent results for *w* parameters above the 6000 m threshold; on the contrary, for *w* values below 5807 m the fine fraction is not significant. These results suggest that *w* values equal to or greater than the range of the dependent variable produce efficient and accurate models, and a marginally better model is obtained for a *w* value exactly equal to the range. Values of *w* lower than the range should be avoided.

As discussed in Section 3.1, sediment values at clam sampling stations were estimated using Kriging interpolation, and this procedure may have induced non-zero correlation between independent variables and model error. As we are not aware of formal inferential procedures to test such correlation, we propose to assess this relationship by analyzing the correlation between model errors and interpolation errors in the estimation of the local values, where the latter error is represented by the Kriging standard error. Since all the correlations (Table 4) are no greater than 0.5, we conclude that the estimated variables, albeit stochastic, are independently distributed with respect to the model error. While in principle this method constitutes a violation of the regression model assumptions, these empirical results indicate that the use of these variables will not hamper the regression estimates.

We considered testing the regression model on a number of interpolations, based on different variogram models, to ensure the robustness of the model. Given the limited sample size, however, we ere able to obtain only very few acceptable variogarm models, and *t*-tests showed no significant difference among the respective interpolations.

A multiplicative, power specification of the bioaccumulation model is discussed in Section 2.1. We believe a linear, multivariate specification, i.e., (Eqs. (6) and (7)), to be an appropriate test of the relationships between bioaccumulation and environmental variables. Following the bioaccumulation literature (Heikens, Peijnenburg, & Hendriks, 2001; Perceval et al., 2002; Sonesten, 2003), a log–linear ($\log(y) = f(x)$) and a log–log model ($\log(y) = f(\log(x))$) on the SC (sediment contaminant) variable were also tested. Table 5 summarizes the results of the linear, log–linear, and log–log models for PCBdl.

Variations in the pseudo-R^2 are only minor, while fewer variables are significant in the logarithmic models. Neither of the logarithmic specifications shows any appreciable improvement over the linear model, therefore we preferred the latter, for its simplicity.

Fig. 4 presents the spatial distribution of the predicted bioaccumulation of PCBdl (the 23 clam sampling stations are also posted).

Fig. 4 provides some important insights: the spatial distribution of pollutant sediment concentration suggests that the main source of the PCBdl in the lagoon is the industrial area (Porto Marghera), located in the central lagoon. Secondary effects are played by

Table 5
Linear, log–linear, and log–log models for PCBdl

		Intercept	PCBdl in sediment	Total organic carbon	pH	Fine fraction	Weight	Lipid content	pseudo-R^2	Log likelihood
Linear model	Coeff.	2165.97	0.21	−223.47		−7.87			0.54	−174.90
	S.E.	257.7653	0.0327	52.3602		3.9314				
	(t value)		6.41	−4.27		−2.00				
	$p(t)$		0.00	0.00	>0.05	0.06	>0.05	>0.05		
Log–linear model	Coeff.	6.1815	0.0002						0.49	−28.17
	S.E.	0.2573	0.0001							
	(t value)		2.6658							
	$p(t)$		0.0148	>0.05	>0.05	>0.05	>0.05	>0.05		
Log–log model	Coeff.	3.5024	0.5893			−0.0151			0.56	−25.69
	S.E.	1.2345	0.1663			0.0072				
	(t value)		3.5428			−2.0892				
	$p(t)$		0.0022	>0.05	>0.05	0.0504	>0.05	>0.05		

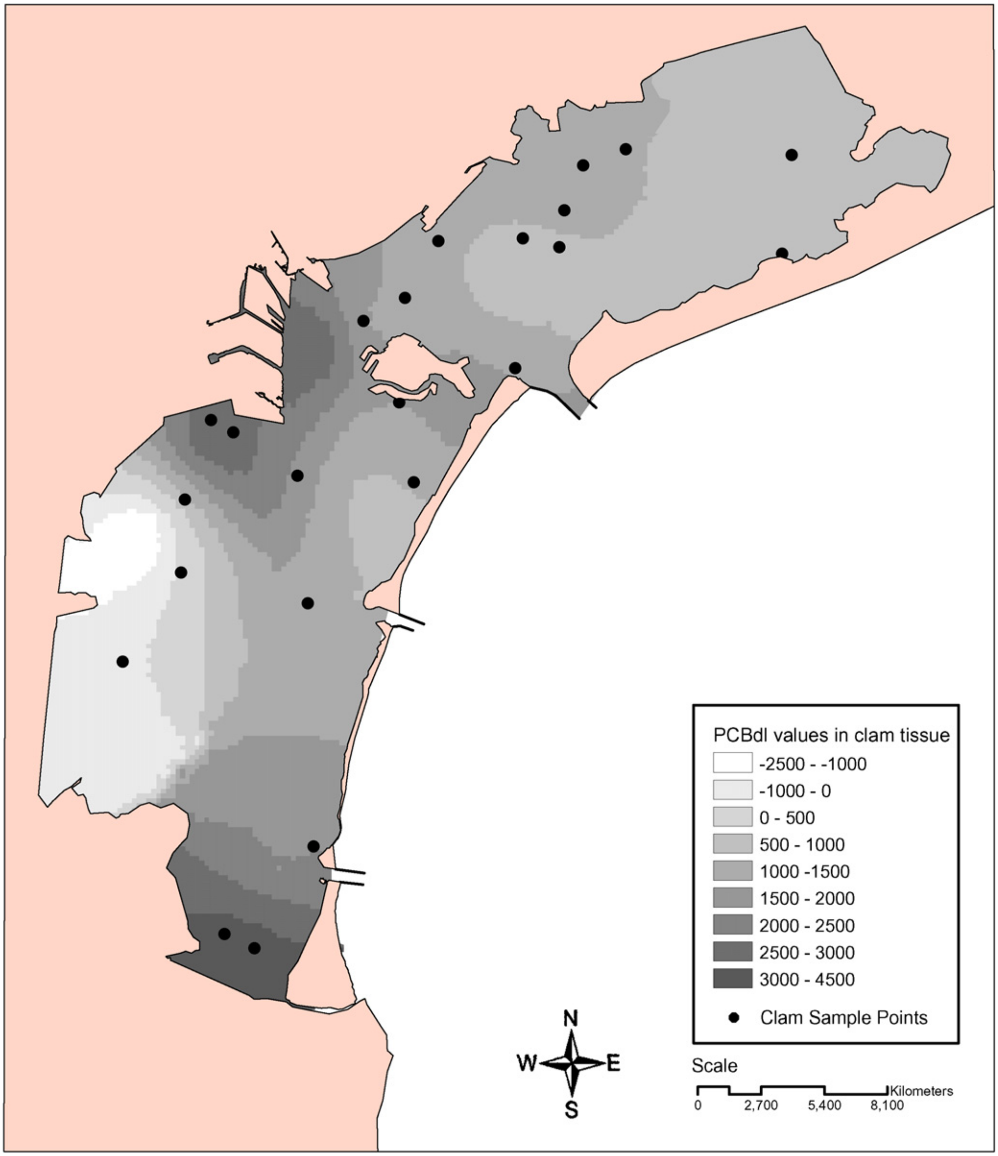

Fig. 4. PCBdl bioaccumulation values predicted by the spatial regression model.

the watershed contributions: two areas appear to be particularly affected: the first is located in the south lagoon, near Chioggia, while the second is located in the north lagoon, corresponding to the Dese and Sile river mouths. Clam bioaccumulation estimated by regression models shows that the PCBdl reaches the highest tissue concentration in organisms located in the south lagoon and in front of the industrial area. It is suitable also to focus monitoring and sampling campaigns in these two areas, thickening the sampling grid. This strategy may be particularly valuable in instances, such as our case study, where data collection and analysis are expensive.

The final step of our method is the use of the coefficients calculated by spatial regression models to characterize exposure at under-sampled locations: we demonstrate this procedure on the PCBdl model. Bioaccumulation of the PCBdl in clam tissue is a positive function of the pollutant concentration in the sediment, and a negative function of total organic content (TOC) and fine fraction (ff).

$$\mathrm{PCBdl_c} = 2165.97 + 0.21\mathrm{PCBdl_s} - 223.47\mathrm{TOC} - 7.87\mathrm{ff} \tag{8}$$

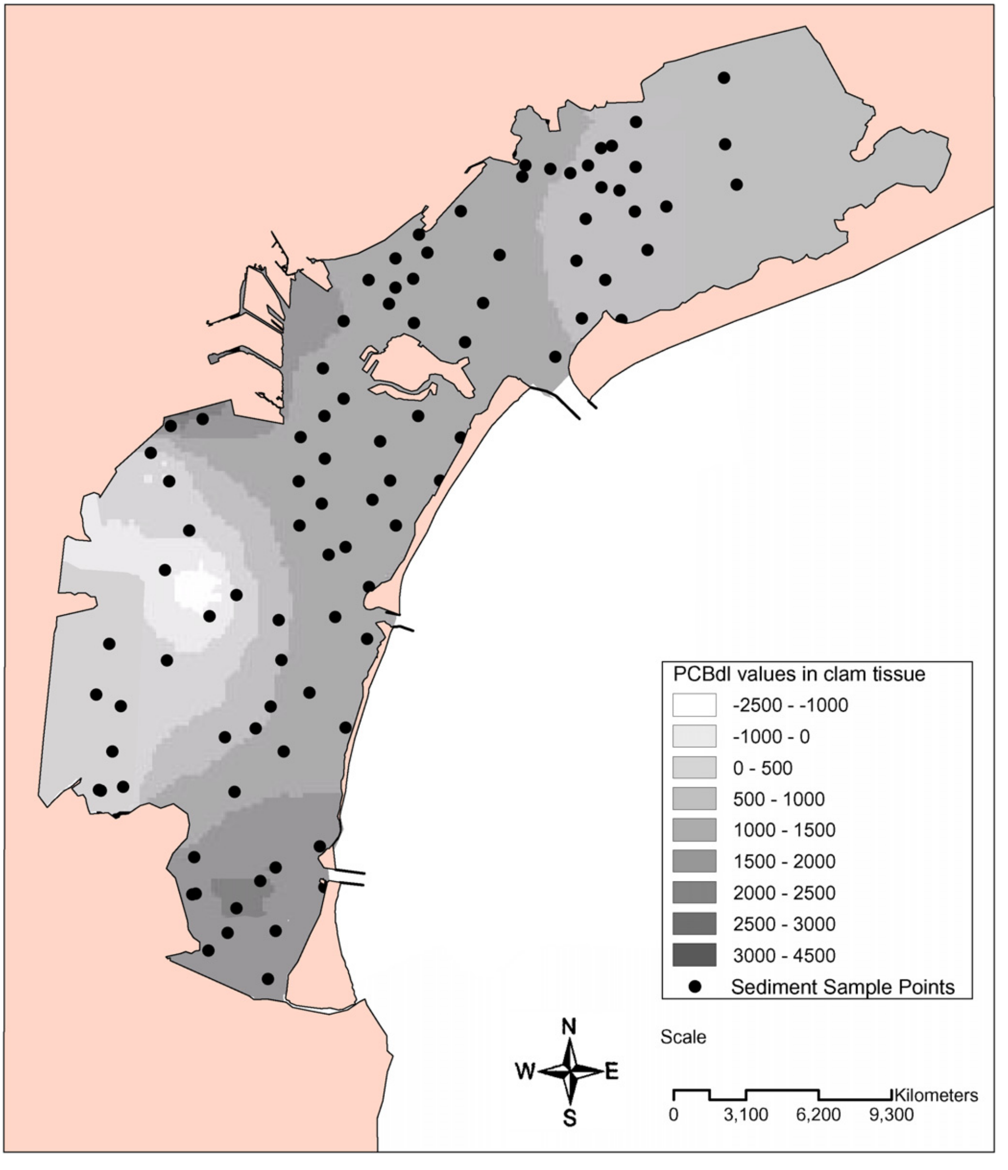

Fig. 5. Coefficient-based estimation of PCBdl bioaccumulation values.

The coefficients in Eq. (8) are used estimate exposure values at locations where only sediment variables were sampled: this allows us to estimate exposure values directly from the model (as opposed to interpolating values) for 95 sample points, since only sediment data are necessary for this calculation. Fig. 5 shows PCBdl bioaccumulation values calculated from Eq. (8) on sediment sample points (95 sediment sampling stations are also posted).

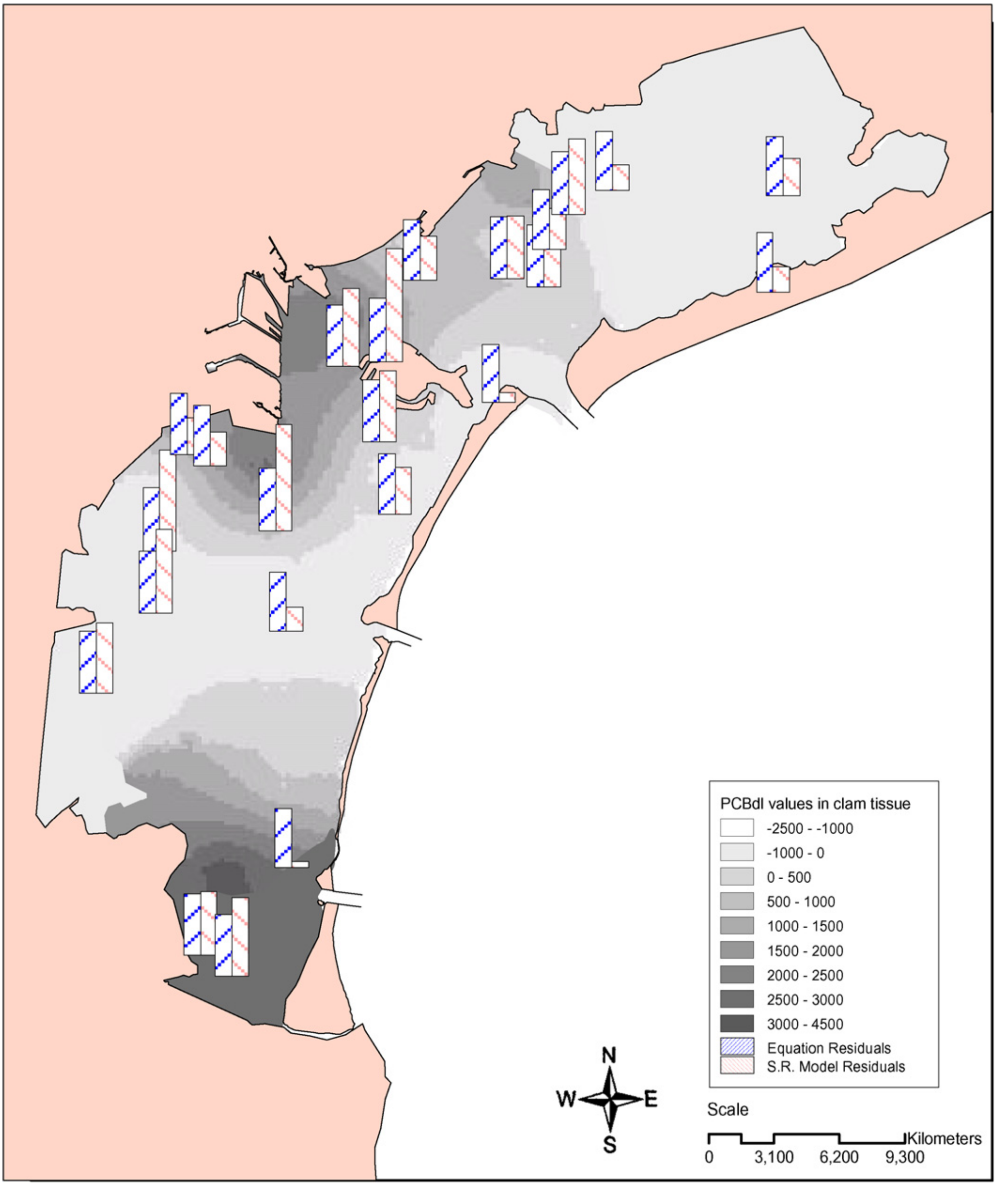

Fig. 6. Residuals of spatial regression model and of coefficient-based estimation.

Due to the different spatial samples for the two sets of variables, we are not able to provide a quantitative assessment of the difference between the predictions in Figs. 4 and 5: map calculations that compute the difference between the two surfaces are based on interpolated values, and statistical tests such as t or χ^2 tests would not account for the location of values; therefore, we only present, in Fig. 6, a comparison of the residuals of the spatial regression model and the residuals of the coefficient-based estimation at the 23 clam sample points -sediment values are, however, estimated (Section 1.1). In the background is a Kriging interpolation of the observed data.

A comparison between the clam data surface, in the background of Fig. 6, and the coefficient-based surface, derived only from sediment data, in Fig. 5, highlights the main discrepancies between the spatial distributions of clams and sediment variables.

Figs. 4–6 present Kriging interpolations of spatial regression predictions, coefficient-based predictions, and clam data, respectively: all these surfaces are represented using a consistent classification, to facilitate their comparison.

4. Discussion

The contributions of our research can be summarized in two main directions. The first is the demonstration of the application of spatial analysis in ERA as a more adequate tool than standard methods; the experiment succeeds in producing meaningful and reliable models that link organism bioaccumulation (receptor) not only to contaminants, but also to the properties of the medium (sediment) where the bioaccumulation process takes place. The second is the development of an integrated procedure to address the inadequacies of sampling for multiple variables; the main result is the provision of a complete protocol to estimate spatial parameters and characterize exposure at under-sampled locations; the procedure is feasible within the assumptions of the spatial regression and geostatistical methods.

The choice of spatial regression models is supported by evidence of spatial autocorrelation in most dependent and independent variables. Comparing standard and spatial regression models, we cannot prove conclusively the superiority of spatial regression models for the sample data: for example, standard errors of the regression coefficients and residual standard errors do not consistently indicate a significant difference between the variance of spatial models and standard models. We believe that such lack of consistency may be largely attributed to the small size of our samples. An important result emerges from the backwards model selection process: significant variables in spatial regression models tend to be much more consistent throughout the various models, and coherent with theoretical expectations, while erratic and less interpretable sets of variables result significant in standard regression models. Additionally, environmental variables (TOC, ff, and pH) tend to be more significant in spatial than in standard regression models. For example, the significant variables in the spatial regression for PCBdl are SC, TOC, and ff, with negative coefficients for TOC and ff and the largest (positive) t value for SC (Table 3). The significant variables for the standard regression model are SC, ff, W (weight), and L (lipid fraction), but both W and L present negative signs, large coefficient (in absolute value), and associated large standard errors.

We can conclude that standard models tend to emphasize individual characteristics, and it is the variability of such values across individuals that inflates coefficient standard errors. This uncertainty in the relationship results in meaningless coefficients and signs, such as those suggesting that bioaccumulation is greater in smaller and leaner individuals. On the contrary, spatial regression models emphasize the role of contaminants and other characteristics of the environment (sediment) where the bioaccumulation process takes place. The spatially autoregressive component of the model reduces the sample variability by means of the spatial weights: such adjustment is produced on a spatial basis, which allows the model to capture the effect of environmental processes interacting with bioaccumulation. Such spatial models present lower coefficient standard errors and represent not only efficient (reliable) but also more effective predictive tools, capable of estimating exposure solely from environmental factors. Thanks to these features, these spatial regression models can be employed to overcome sampling gaps (as in our case study), and, perhaps more importantly, to identify areas of potential ecological concern for other species exposed to the same factors.

The most meaningful and satisfactory spatial regression models are those for organic micropollutants (Section 3.2). Spatial regression models for inorganic micropollutants (metals) can only be viewed as a preliminary step toward site-specific analyses, but their poor performance suggests that additional variables may be necessary to explain the effective bioavailability of metals. Bioavailability is affected by sediment characteristics such as concentration of iron, manganese and aluminum oxides (FeO_x, MnO_x, AlO_x) (Bendell-Young & Harvey, 1991; Bendell-Young, Chouinard, & Pick, 1994; Bryan & Langston, 1992; Janssen et al., 1997), and acid volatile sulphide (Chapman, Wang, Janssen, Persoone, & Allen, 1998), all of which should be included as independent variables.

Non-coincident sampling of multiple variables at non-coincident locations is a common occurrence in applied multivariate analysis; the difficulties induced in multiple regression are comparable to the case of samples drawn from spatial units that do not overlap exactly, more frequent in the social sciences (e.g., variables sampled at postal codes and census enumeration areas). We address this issue by integrating an optimal spatial interpolator, i.e., Kriging, to estimate variable values at un-sampled locations that will be used in the spatial regression model; we suggest complementing this procedure with a test to assess the correlation between interpolation error and regression error, since the assumption of a non-stochastic independent variable cannot be met. The entire integrated procedure can represent an experimental protocol for exposure characterization, as well as for other environmental and socio-economic applications. It is important to note that this procedure is based on the assumption of stationarity of the spatial process; as discussed in Section 2.2.2 we tested stationarity and spatial dependence, and consequently chose a spatially autoregressive over a geographically weighted model.

This procedure results in a characterization of exposure at under-sampled locations based on the use of coefficients calculated by spatial regression models. Assessing the difference between this method and the regression model is not possible with the available data: therefore, we discuss a visual comparison of Fig. 4 (regression model) and Fig. 5 (coefficient-base estimate). The two maps appear overall consistent, as both identify the main source of PCBdl in the industrial area (Porto Marghera), in the central lagoon; the highest PCBdl values are estimated near the industrial area and in the southern tip of the lagoon. Due to a finer sampling grid, the coefficient-based estimation generally produces more detailed results, e.g., in the south and south-central lagoon; additionally, estimates of the coefficient-based model are less influenced by the values of individual points. Negative signs (TOC and ff) and large coefficient values (TOC) do not seem to affect the estimated values, which appear to be overall equally conservative. More accurate estimates are particularly important at the local level, supporting the value of this model as a planning tool for fishing and fish farming zones. The coefficient-based model highlights some areas of potential concern, such as the south-east edge of the lagoon and the spots just north of the city of Venice: such zones would not emerge from the direct application of the regression model, as they are heavily under-sampled, as shown by the sampling stations posted on the maps. Conversely, lower contamination values in the north and north-east lagoon are confirmed by the coefficient-based estimates, derived from a much finer sampling grid. Due to its higher environmental quality, fishing and fish farming should be promoted in this area.

Figs. 4 and 5 summarize the results of numerous spatial analysis applications. They provide tools to estimate exposure levels at any location, as well as to identify and analyze hot spots. Using these tools, pollutant sources can be identified and monitored, and appropriate sampling and monitoring campaigns can be designed, focusing on the most problematic areas. This strategy may reduce the cost of analysis and interventions, while improving its effectiveness (World Bank, 1997).

5. Conclusion

Using the case study of contaminated sediments in the Venetian lagoon, we demonstrate how geographic information can and should be used to improve traditional models for the estimation of exposure values in Ecological Risk Assessment (ERA). A complete protocol is proposed to integrate geostatistics and multivariate spatial regression analysis to produces efficient and reliable models. The most meaningful models are obtained for organic compounds: organism bioaccumulation (receptor) is linked not only to contaminants, but also to the properties of the medium (sediment) where the bioaccumulation process takes place. Based on the assumption of stationarity of the spatial process, we demonstrate how spatial regression coefficients can also be used to estimate exposure values from environmental variables at under-sampled locations.

The explicit link between the organism and the characteristics of its environment, the reliability of the estimates, and the provision of estimates for under-sampled locations drastically enhance the practical value of spatial models and the feasibility of their use as effective planning tools.

Acknowledgements

We wish to thank the Water Authority of Venice (Magistrato alle Acque) and the "Venezia Nuova Consortium", as well as the Health Authority of the Veneto Region for supplying the data used in this study. We appreciate their work and the assistance they provided with the management and interpretation of the datasets.

We would also like to thank the editor of CEUS and the referees who reviewed our manuscript: their insightful comments helped us to achieve a deeper understanding of our own work and to improve the quality of the research and its presentation.

References

Anselin, L. (1988). *Spatial econometrics: Methods and models*. New York: Kluwer.

Anselin, L. (1993). SpaceStat tutorial. Regional Research Institute. West Virginia University, Morgantown, West Virginia.

Anselin, L. (1998). Exploratory spatial data analysis in a geocomputational environment. In P. A. Longley, S. M. Brooks, R. McDonnell, & B. MacMillan (Eds.), *Geocomputation* (pp. 77–94). Chichester: John Wiley & Sons.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. Upper Saddle River: Prentice Hall.

Bendell-Young, L., Chouinard, J., & Pick, F. R. (1994). Metal concentrations in chironomids in relation to peatland geochemistry. *Archives of Environmental Contamination and Toxicology, 27*, 186–194.

Bendell-Young, L., & Harvey, H. H. (1991). Metal concentration in chironomids in relation to the geochemical characteristics of surficial sediments. *Archives of Environmental Contamination and Toxicology, 21*, 202–211.

Bertazzon, S. (2003). A definition of contiguity for spatial regression analysis in GISc: Conceptual and computational aspects of spatial dependence. *Rivista Geografica Italiana, 2*, 247–280.

Bertazzon, S., Carlon, C., Critto, A., Marcomini, A., & Zanetto, G. (2000). Integration of spatial analysis and ecological risk assessment in a GIS environment. The case study of the Venetian lagoon contaminated sediments. In *Proceedings of GIS/EM4, 4th conference on GIS and environmental modeling, Banff, Canada, Vol. 122*.

Bryan, G. W., & Langston, W. J. (1992). Bioavailability, accumulation and effects of heavy metals in sediments with special reference to United Kingdom estuaries: A review. *Environmental Pollution, 76*, 89–131.

Chapman, M. P., Wang, F., Janssen, C., Persoone, G., & Allen, E. H. (1998). Ecotoxicology of metals in aquatic sediments: Binding and release, bioavailability, risk assessment and remediation. *Canadian Journal of Fisheries and Aquatic Sciences, 55*, 2221–2243.

Davis, J. (2002). *Statistics and data analysis in geology* (3rd ed.). New York: John Wiley & Sons.

Facca, C., Sfriso, A., & Socal, G. (2002). Changes in abundance and composition of phytoplankton and macrophytobenthos due to increased sediment fluxes in the Venice lagoon, Italy. *Estuarine, Coastal and Shelf Science, 54*, 773–792.

Fotheringham, A. S., Brundson, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester: Wiley.

Gamma Design Software (2001). Available from http://www.gammadesign.com/.

Gatrell, A. C. (1983). *Distance and space: A geographical perspective*. Oxford: Clarendon Press.

Getis, A., & Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis, 36*, 90–104.

Gobas, A. P. C. F. (1993). A model for predicting the bioaccumulation of hydrophobic organic chemicals in aquatic food-webs: Application to Lake Ontario. *Ecological Modelling, 69*, 1–17.

Griffin, R. A. (1991). Movement of PCBs and other persistent compounds through soil. *Water Science and Technology, 13*, 1153–1163.

Grondona, M. O., & Cressie, N. (1991). Using spatial considerations in the analysis of experiments. *Technometrics, 33*, 381–392.

Guerzoni, S., Rossini, P., Molinaroli, E., Rampazzo, G., & Raccanelli, S. (2004). Measurements of atmospheric deposition of polychlorinated dibenzo-*p*-dioxins and dibenzofurans in the lagoon of Venice, Italy. *Chemosphere, 54*, 1309–1317.

Gujarati, D. (2003). *Basic econometrics* (4th ed.). New York: McGraw-Hill.

Haining, R. (1997). *Spatial data analysis in the social and environmental sciences*. Cambridge: Cambridge University Press.

Heikens, A., Peijnenburg, W. J. G. M., & Hendriks, A. J. (2001). Bioaccumulation of heavy metals in terrestrial invertebrates. *Environmental Pollution, 113*, 385–393.

Im, S. H., Strause, K. D., Giesy, J. P., Chang, Y. S., Matsuda, M., & Wakimoto, T. (2004). Concentrations and accumulation profiles of polychlorinated dibenzo-*p*-dioxins and dibenzofurans in aquatic tissues, and ambient air from South Korea. *Chemosphere, 55*, 1293–1302.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. Oxford: University Press.

Janssen, R. P. T., Posthuma, L., Baerselman, R., den Holander, H. A., van Veen, R. P. M., & Peijnenburg, W. J. G. M. (1997). Equilibrium partitioning of heavy metals in Dutch soils. II. Prediction of metal accumulation in earthworms. *Environmental Toxicology and Chemistry, 16*, 2479–2488.

Jarvinen, A. W., Nordling, B. R., & Henry, M. E. (1983). Chronic toxicity of dursban to the fathead minnow and the resultant acetylcholinesterase inhibition. *Ecotoxicology and Environmental Safety, 7*, 423–434.

Jarvis, C. H., Stuart, N., Baker, N., & Morgan, R. (1999). To interpolate and thence to model, or vice versa? In B. Gittings (Ed.). *Integrating information infrastructures with GI technology: Innovations in GIS* (Vol. 6, pp. 229–242). London: Taylor & Francis.

Juang, K. W., & Lee, D. Y. (1998). A comparison of three Kriging methods using auxiliary variables in heavy metals contaminated soils. *Journal of Environmental Quality, 27*, 355–363.

Legendre, P., Dale, M. R. T., Fortin, M.-J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography, 25*, 601–615.

Leonte, D., & Schofield, N. (1996). Evaluation of a soil contaminated site and clean up criteria: A geostatistical approach. In R. M. Srivastava, S. Rphuani, M. V. Cromer, A. I. Johnson, & A. J. Desbarats (Eds.), *Geostatistics for environmental and geotechnical applications*. American Society for Testing and Materials.

Linkov, I., Burmistrov, D., Cura, J., & Bridges, T. (2002). Risk-based management of contaminated sediments: consideration of spatial and temporal patterns in exposure modeling. *Environmental Science and Technology, 36*, 238–246.

Mahony, D. J., Di Toro, M. D., Gonzalez, M. A., Curto, M., Dilg, M., & De Rosa, D. L. (1996). Partitioning of metals to sediment organic carbon. *Environmental Toxicology and Chemistry, 15*, 2187–2197.

Mendez, E., Giudice, H., Pereira, A., Inocente, G., & Medina, D. (2001). Total mercury content and fish weight relationship in swordfish (*Xiphias gladius*) caught in the southwest Atlantic Ocean. *Journal of Food Composition and Analysis, 14*, 453–460.

Montanes, J. F. C., Van Hattum, B., & Deneer, J. (1995). Bioconcentration of chlorpyrifos by the freshwater isopod *Asellus aquaticus* in outdoor experimental ditches. *Environmental Pollution, 88*, 137–146.

Mountouris, A., Voutsas, E., & Tassios, D. (2002). Bioconcentration of heavy metals in aquatic environments: The importance of bioavailability. *Marine Pollution Bulletin, 44*, 1136–1141.

Munns Jr., W. M., Black, D. E., Gleason, T. R., Salomon, K., Bengtson, D. A., & Gutjahr-Gobell, R. (1997). Evaluation of the effects of dioxin and PCBs on *Fundulus heteroclitus* populations using a modeling approach. *Environmental Toxicology and Chemistry, 16*, 1074–1081.

Nathanail, C. P., Ferguson, C. C., Brown, M. J., & Hooker P. J. (1998). A geostatistical approach to spatial risk assessment of lead in urban soils to assist planners. In *Proceedings of the 8th international congress IAEG, Vancouver, Canada*, pp. 369–376.

Perceval, O., Pinel-Alloul, B., Méthot, G., Couillard, Y., Giguére, A., Campbell, P. G. C., et al. (2002). Cadmium accumulation and metallothionenin synthesis in freshwater bivalves (*Pyganodon grandis*): Relative influence of the metal exposure gradient versus limnological variability. *Environmental Pollution, 118*, 5–17.

Serrano, R., Hernandez, F., Pena, J. B., Dosda, V., & Canales, J. (1995). Toxicity and bioconcentration of selected organophosphorus pesticides in *Mytilus galloprovincialis* and *Venus gallina*. *Archives of Environmental Contamination and Toxicology, 29*, 284–290.

Sonesten, L. (2003). Fish mercury levels in lakes—Adjusting for Hg and fish-size covariation. *Environmental Pollution, 125*, 255–265.

Tobler, W. R. (1979). Cellular geography. In S. Gale & G. Olsson (Eds.), *Philosophy in Geography* (pp. 379–386). Dordecht: Reidel.

US-EPA (1992). Framework for ecological risk assessment. Risk Assessment Forum, EPA/630/R-92/001. US Environmental Protection Agency, Washington, DC.

US-EPA (1998). Guidelines for ecological risk assessment. Risk Assessment Forum, EPA/630/R-95/002F. US Environmental Protection Agency, Washington, DC.

US-EPA (2000). Ecological soil screening level guidance. Office of Emergency and Remedial Response. Draft Report. US Environmental Protection Agency, Washington, DC.

Van den Berg, M., Birnbaum, L., Bosveld, A. T. C., Brunstorom, B., Cook, P., Freely, M., et al. (1998). Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environmental Health Perspective, 106*, 775–792.

World Bank report (June 27, 1997). Brazil: Managing pollution problems. The Brown Environmental Agenda. Report #16513-BR.