

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Using a Multi-Scale Geostatistical Method for the Source Identification of Heavy Metals in Soils

Nikos Nanos<sup>1</sup> and José Antonio Rodríguez Martín<sup>2</sup>

<sup>1</sup>*School of Forest Engineering - Madrid Technical University  
Ciudad Universitaria s/n, Madrid*

<sup>2</sup>*I.N.I.A. Department of the Environment, Madrid  
Spain*

## 1. Introduction

For quite a long time, soil has been considered a means with a practically unlimited capacity to accumulate pollutants without immediately producing harmful effects for the environment or for human health. Presently, however, we know that this is not true. Public awareness has been raised on the harmful potential of some soil trace elements –commonly known as heavy metals- that can accumulate in crops and may end up in human diet through the food chain. Many studies have confirmed that heavy metals may accumulate and damage crops or even mankind (Otte et al., 1993; Dudka et al., 1994; Söderström, 1998). Along these lines, the most dangerous metals owing to their toxicity for human beings are Cd, Hg and Pb (Chojnacka et al., 2005).

Natural concentration of heavy metals in soil is generally very low and tends to remain within very narrow limits to ensure an optimum ecological equilibrium. Nonetheless, human activities that involve emitting large quantities of heavy metals into the environment have dramatically increased natural concentrations in the last century. Although soils are quite capable of cushioning anthropogenic inputs of toxic substances, there are times when this capacity is exceeded, which is when a pollution problem arises.

The natural concentration of heavy metals in soils depends primarily on geological parent material composition (Tiller, 1989; Ross, 1994; Alloway, 1995; De Temmerman et al., 2003; Rodríguez Martín et al., 2006). The chemical composition of parent material and weathering processes naturally conditions the concentration of different heavy metals in soils (Tiller, 1989; Ross, 1994). In principle, these heavy metals constitute the trace elements found in the minerals of igneous rocks at the time they crystallize. In sedimentary rocks, formed by the compactation and compression of rocky fragments, primary or secondary minerals like clays or chemical precipitates like CaCO<sub>3</sub>, the quantity of these trace elements depends on the properties of the sedimented material, the matrix and the concentrations of metals in water when sediments were deposited. In general, concentrations of heavy metals are much higher in igneous rocks (Alloway, 1995; Ross, 1994). Nonetheless, these ranges vary widely, which implies that the natural concentration of heavy metals in soil will also vary widely.

Variability of heavy metals in soil is also associated with the variability of the physico-chemical properties of soil. pH, organic matter, clay minerals, metal oxides, oxidation-reduction reactions, ionic exchange processes, or adsorption, desorption and complexation phenomena, are the main edaphic characteristics relating to retention of metals in soil. Nonetheless, some metals show a strong affinity to organic matter, while others have a strong affinity to clays, and to Fe and Mn oxides; yet they are elements which tend to precipitate in a carbonate form. Most of these soil properties will depend on the geological parent material. However, they will also be subject to not only the electric charge of metals in relation to the former saturation of other ions, but to all the interactions taking place naturally among the various edaphic parameters.

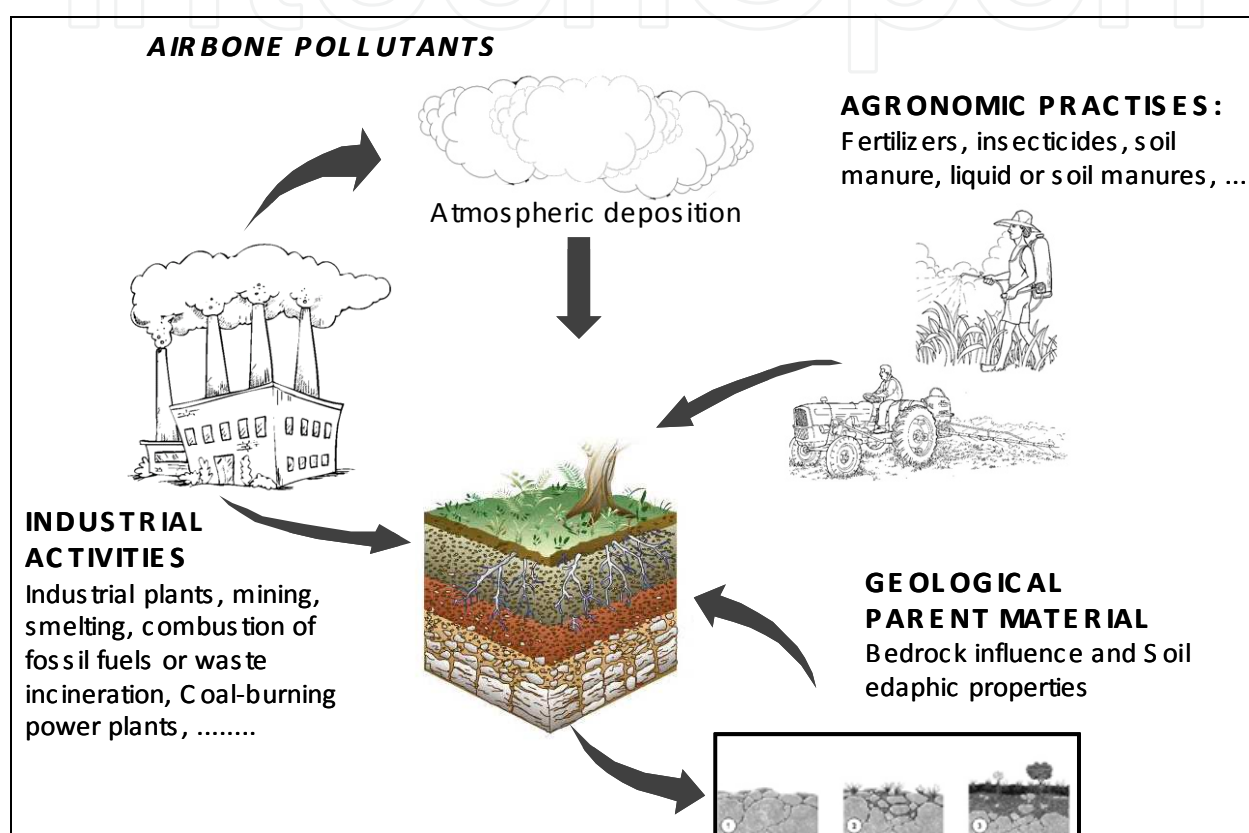


Fig. 1. Heavy metal input in soil

### 1.1 Anthropogenic influence on the content and distribution of heavy metals in soil

Many productive activities like mining, smelting, industrial activity, power production, pesticides production or waste treatment and spillage, represent sources of metals in the environment (Gzyl, 1999; Weber & Karczewska, 2004). The concentration of metals in soil can increase directly, or be due to the impact of atmospheric deposition (Figure 1) caused by proximity to industrial plants (Colgan et al., 2003) or through fossil fuel combustion (Sanchidrian & Mariño, 1980; Martin & Kaplan, 1998). Thus the atmosphere, which is an important means of transport for heavy metals originating from various emission sources, is the first factor that enriches soil in heavy metals. Although the natural entrance of metals to the atmosphere derives from volcanoes and the evaporation of the earth's crust and oceans, the main current input is of an anthropogenic origin. Airborne pollutants produced by mining, combustion of fossil fuels or waste incineration have significantly increased the emission of

heavy metals into the atmosphere, which are subsequently deposited on the soil surface by deposition (Figure 2). The effects of this pollution can become evident in soil at a distance of hundreds of kilometers from the emission source (Hutton, 1982; Nriagu, 1990; Alloway & Jackson, 1991; Navarro et al., 1993; Engle et al., 2005).

In addition, normal agricultural practices may cause enrichment of heavy metals (Errecalde et al., 1991; Kashem & Singh, 2001; Mantovi et al., 2003). These practices are an important source of Zn, Cu, and Cd (Nicholson et al., 2003) due to the application of either liquid and soil manure (or their derivatives, compost or sludge) or inorganic fertilizers. Finally, there is also the pollution resulting from spillages, which are normally isolated and easy to identify, generally because they lead to very high values which multiply the expected soil content by several units. These sporadic sources of pollution are not usually observed in areas employed for growing crops.

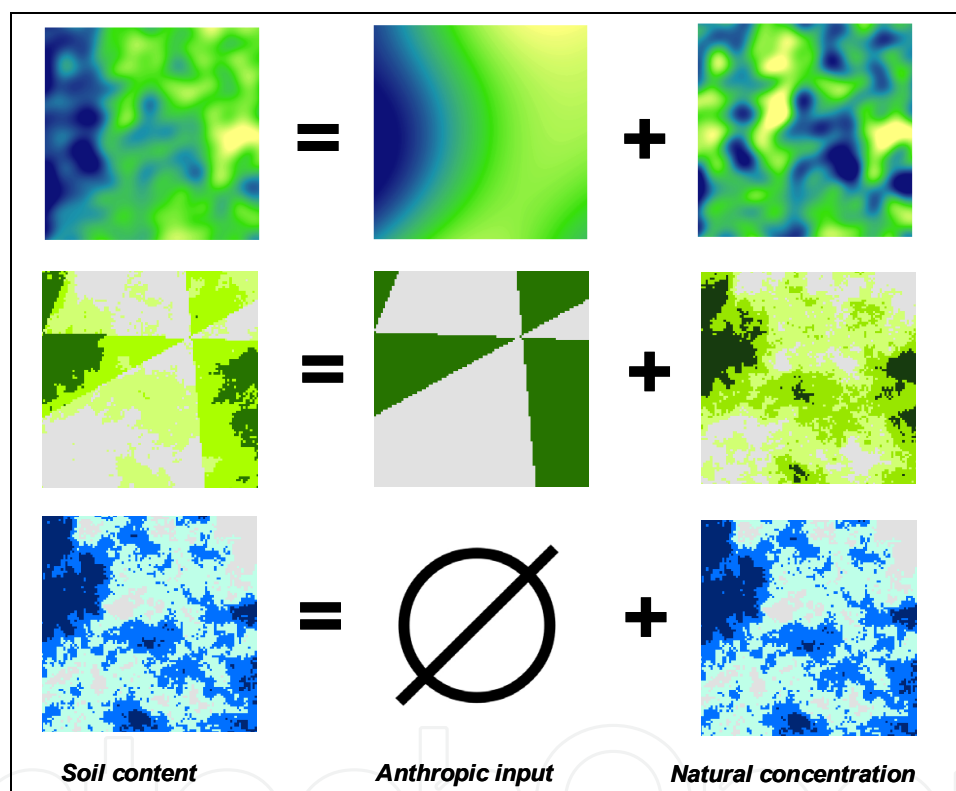


Fig. 2. Schematic representation of multiple sources of heavy metals in soil. In the upper panels, we present the spatial distribution of a heavy metal enriched by airborne pollution – and subsequent deposition-. In the middle panels, we provide a hypothetical example of a heavy metal enriched by agricultural practices such as fertilization (green indicates input of heavy metals). In the lower panels, natural soil content is assumed to not be enriched by any human activity. Note that in real case studies we may only observe the sum of two inputs.

## 1.2 Identification of sources of soil heavy metals

Source identification and apportionment of metal elements in soil are not straightforward. Quantities of metals introduced into soil through industrial activities, or any other human activity, do not provide any trace of their anthropogenic origin and, once inside soil, they behave like any other similar natural analogue which is already in soil. Likewise metal like

copper, which forms part of soil after fertilizing farmland, shows no distinguishing element of its origin. This makes us wrongly think that human input does not exist if the total concentration in soil does not reach levels considered to be polluting. Consequently, what we observe in the analysis of these elements in soil tends to result from summing the two inputs (Figure 2).

Despite the difficulty of separating human input from natural input, the separation task is greatly needed for correct edaphic resource management and to prevent its pollution. In this chapter, we present a multiscale geostatistical method known as a factorial kriging analysis which -under certain circumstances- can provide a mathematical framework to distinguish natural soil enrichment from that of an anthropogenic kind in heavy metals. The method was initially presented by Matheron (1982) and has been used repeatedly in soil science (Goovaerts, 1992, Castrignano, 2000, Rodriguez et al., 2008).

### 1.3 Description of the study areas used in the analysis

The information used in this chapter originates from the sampling and the analysis of the two most important hydrographic basins in Spain (Figure 3) in a study conducted to obtain reference values in Spanish soils (Rodríguez et al., 2009). They all present different lithologies and geologies, as well as common edaphic processes. Likewise, they also reflect possible inputs from both farming treatments and industrial activities which modify the contents and distribution of these elements in each valley.



Fig. 3. Localization of the Duero and Ebro basins. Soil samples are plotted over the two river basins

### 1.3.1 The Duero basin

The Duero river basin is the largest of its kind in the Iberian Peninsula, with a total surface of 97,290 km<sup>2</sup>, of which 78,954 km<sup>2</sup> are found in Spanish territory (Confederación Hidrográfica del Duero, [www.chduero.es](http://www.chduero.es)). From a geological point of view, the hydrographic Duero basin consists of a well-defined geological unit, the Duero depression and its borders. The Duero Basin is an intraplate continental basin which developed from the Late Cretaceous to the Late Cenozoic. Over this time span, the basin acted as a foreland basin of the surrounding Cantabrian Zone and the Basque-Cantabrian Range in the north, the Iberian Range in the east and the Central System in the south, which constitute complex fold and thrust belts. These Alpine compressional ranges, constituted by Palaeozoic and Mesozoic rocks, are thrusting, in many cases, the Cenozoic deposits of the Duero Basin, which are virtually undeformed (Gomez et al., 2006). The basin is filled mainly with siliciclastic sediments on the margins and evaporites in central areas, showing an endorheic arrangement (Tejero et al., 2006). The basin's general facies distribution corresponds to a continental foreland basin model with alluvial fan deposits grading into alluvial plains, and evaporitic and carbonated lacustrine environments toward the centre of the basin (Gomez et al., 2006).

The total population in the basin is around 2,210,541 inhabitants (Municipal Register, 2006), which has barely varied in the last hundred years. The Spanish Autonomous Community of Castilla y León is one of the main cereal-growing areas in Spain. Apart from pulses, such as carob and chickpea, the practice of growing sunflowers has extended in the southern countryside. However, the number of cultivated vineyard hectares (56,337 ha) lowered considerably in the last three decades of the past century.

### 1.3.2 The Ebro basin

The Ebro Valley, in the northeast region of the Iberian Peninsula, is framed by three mountain ranges: the Pyrenees to the north, the Iberian Chain to the southwest, and the Catalanian Coastal Ranges to the southeast. The structural development of these ranges controlled the evolution of this basin in tectonic and structural terms, and as regards stratigraphic and sedimentologic aspects. The Pyrenean range is a fold and thrust belt that developed during the Tertiary (the following 60 million years) as the Iberian block converged toward the European Plate. Its metamorphic core marks the border between France and Spain, with foreland structures verging into both countries. The tertiary stratigraphic units also increase in thickness northwardly. The tectonic load of the allochthonous units and the formation of a cold lithospheric root during Pyrenean shortening induced the deflection of the Ebro basin (Brunet, 1986, Roure et al., 1989). The depth of the Tertiary basin increases northwardly, reaching values of 4000 m beneath the sea level below the Pyrenees (Riba et al., 1983). The most exposed rocks within the basin area are of the Oligocene-Miocene age (including clastic, evaporite and carbonate facies) and of an alluvial and lacustrine origin (Riba et al., 1983; Simon-Gomez, 1989). During the Quaternary, incision of the drainage system caused the isolation of structural platforms, the tops of which are composed of near-horizontal Neogene limestones. Contemporaneously, several nested levels of alluvial terraces and sediments developed (Simón and Soriano, 1986). The Quaternary levels comprise mainly gravels, sand and slits.

The Ebro river region, with a population of around 3.25 million, is intensively industrialized. The Ebro river region is also an important agricultural area in Spain, with 4.2 million ha (Fig. 1) of agricultural topsoil (the total basin area is 9.5 million ha).

### 1.3.3 Soil samples and chemical analysis

The sampling scheme was based on an 8x8 km grid. Samples were located using GPS and topographic maps on a scale of 1:25,000. Each sample was defined as a composite made up of 21 subsamples collected with the Eijkelpamp soil sampling kit from the upper 25 cm of soil in a cross pattern. Further details can be found in Rodríguez et al. (2009).

Soil samples were air-dried and sieved with a 2 mm grid sieve. Soil texture was determined for each sample. After shaking with a dispersing agent, sand (2 mm–63  $\mu$ m) was separated from clay and silt with a 63  $\mu$ m sieve (wet sieving). A standard soil analysis was carried out to determine the soil reaction (pH) in a 1:2.5 soil-water suspension (measured by a glass electrode CRISON model Microph 2002) and organic matter (%) by dry combustion (LECO mod. HCN-600) after ignition at 1050°C and discounting the carbon contained in carbonates. Carbonate concentration was analyzed by a manometric measurement of the CO<sub>2</sub> released following acid (HCl) dissolution (Houba et al., 1995).

Metal contents (Cr, Ni, Pb, Cu, Zn, Hg and Cd) were extracted by aqua regia digestion of the soil fraction in a microwave (ISO 11466, 1995). Heavy metals in soil extracts were determined by optical emission spectrometry (IPC) with a plasma spectrometer ICAP-AES. Mercury in soil extracts was determined by cold vapor atomic absorption spectrometry (CVAAS) in a flow-injection system. The summary statistics of soil parameters and heavy metal contents are listed in Table 1.

	Duero basin 721 soil samples					Ebro basin 624 soil samples				
	Mean	Median	S.D.	1st Qu	3rd Qu	Mean	Median	S.D.	1st Qu	3rd Qu
<i>S.O.M</i>	1.74	1.3	1.494	0.88	2.08	2.24	2	1.41	1.4	2.6
<i>Soil pH</i>	7.19	7.7	1.234	6	8.3	8.04	8	0.59	8	8.4
<i>E.C</i>	0.18	0.15	0.16	0.1	0.21	0.59	0.27	0.85	0.21	0.45
<i>CaCO<sub>3</sub></i>	9.41	3	14.45	1	10.1	29.68	31	16.08	20	40
<i>Sand</i>	59.2	61	19	46	75	38.67	38	17.09	26	50
<i>Silt</i>	20.9	19	12.7	10	31	22.05	21	8.7	16	27
<i>Clay</i>	19.9	19	10.6	12	26	39.41	39.4	13.08	30	48
<i>Cr</i>	20.53	18	14.9	10	27	19.82	18	11.18	13	24
<i>Ni</i>	15.08	13	9.99	8	20	19.26	18	8.64	13	23
<i>Pb</i>	14.06	13	6.79	9	17	16.98	15	7.63	12	20
<i>Cu</i>	11.01	10	7.84	6	15	16.68	13	11.89	10	21
<i>Zn</i>	42.42	38	23.01	27	53	57.4	55	24.27	41	69
<i>Hg</i>	42.05	30	58.43	18	50	33.83	27	38.66	16	44
<i>Cd</i>	0.159	0.1	0.14	0.07	0.2	0.413	0.4	0.159	0.3	0.5

1st Qu, 3rd Qu, first and third quartile; SD, standard deviation ; SOM, soil organic matter (%); CaCO<sub>3</sub>, carbonates (%); EC, Soil electrical conductivity (dS m<sup>-1</sup>)

Table 1. Statistical summary of metal concentrations of soil (in mg/kg for Zn, Fe, Cu, Cd and  $\mu$ /kg for Hg) and some soil properties.

## 2. Factorial kriging analysis

### 2.1 Univariate analysis

Usually physical and chemical soil variables are intrinsically structured around more than one scale of variation, and it is up to the researcher to identify the most important one(s) for his/her study. In factorial kriging -and in geostatistics in general- identification of the scale (or scales) of variation of a variable is done via the variable's sample (or experimental) variogram.

Let us consider a regionalized variable sampled at  $N$  points  $\{x_a : a = 1, \dots, N\}$  such as the Cd concentration in Ebro basin top soil. The sample variogram for Cd is then computed according to the formula:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{x_\beta - x_\alpha \approx h} [z(x_\beta) - z(x_\alpha)]^2 \quad (1)$$

where  $N(h)$  is the number of pairs of data locations separated by, approximately, distance  $h$  and  $\hat{\gamma}(h)$  the semivariance for lag  $h$ . In fact, the sample variogram is constructed by grouping pairs of observations into several discrete distance classes or lags. The average separation distance of all the pairwise points falling within lag  $h$  is plotted against half the average semivariance for each lag computed with formula 1, giving raise to a scatter plot similar to that depicted in Figure 4. Typically, semivariance exhibits an ascending behavior near the origin ( $h=0$ ), whereas at longer separation distances, it levels off at a maximum value called the variogram sill. The distance at which the sill is reached is called the variogram range, while the term nugget is used for the semivariance value at a distance of  $h=0$ .

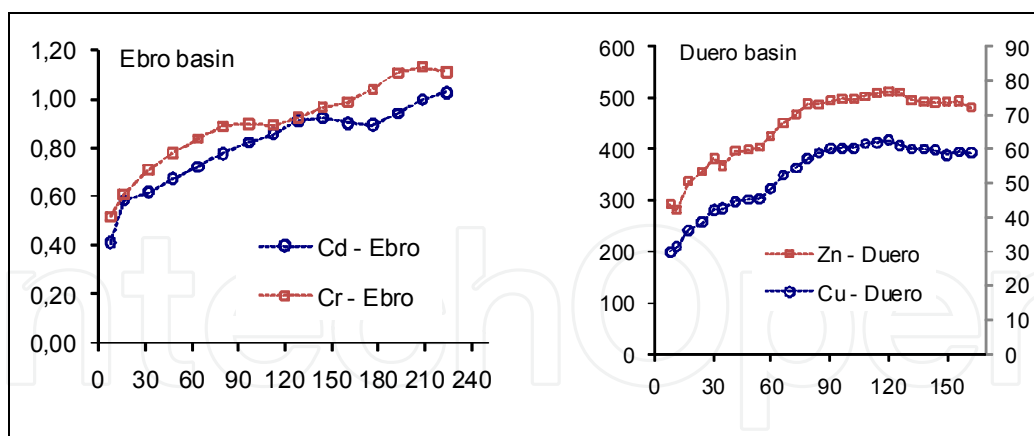


Fig. 4. The sample variograms for Cd and Cr (Ebro river basin) and for Zn and Cu (Duero basin). Horizontal distance is measured in km; vertical axes project the semivariance for a certain distance lag. Note also that the semivariance values for the Ebro have been standardized to unit variance (this is not the case for Duero basin). Note too how the slope of the sample variograms changes before reaching the sill, indicating possible variation on several spatial scales.

Sample variograms, like those shown in Figure 4, indicate that the variable is not distributed randomly in space, rather it is spatially correlated -with an extended spatial correlation



equal to the range of the variogram- so that the pairs of observations separated by a distance shorter than the range of the variogram are more similar on average than usual. When sample variograms also show two or more distinct slopes while ascending toward the sill, the variable can be considered a multi-scale distributed variable. In the Ebro case study, for instance, the sample variograms for Cd and Cr exhibit two different slopes before the sill. More specifically, Cd presents a steep slope in the first distance lag (20 km), a second one up to approximately 100 km, and a last slope change until ca. 220 km. Cr, on the other hand, exhibits two slopes at approximately 90 km and 200 km. Conversely, the sample variograms for Zn and Cu in the Duero basin seem to reach the variogram sill at distance of 100 km. Zinc presents a considerable step ascension in its semivariance at short distances of less than 20 km.

### 2.1.1 Variogram modeling in the univariate case

After the sample variogram has been calculated, it is modeled using either a unique variogram model or a combination of more than one variogram models (also called structures). The variables presenting spatial variation on a unique spatial scale are conveniently modeled using a combination of two variogram models: a nugget model and a structure with a range of spatial correlation equal to the range of the sample variogram. However, whenever multiscale variation is present in the sample variogram, models with more structures should be adopted for modeling. Several variogram models can be used at this stage (a detailed list of variogram models and their characteristics can be found in Chilés and Delfiner (1999)). However, for multiscale variation, modeling practice has shown that spherical models are the most convenient. If we consider that  $k=1, \dots, q$  denotes the number of structures used to model the experimental variogram, so the variogram model (also called the linear model of regionalization) can be written as:

$$\gamma(\mathbf{h}) = \sum_{k=1}^q \gamma_k(\mathbf{h}) = \sum_{k=1}^q b_k g_k(\mathbf{h}) \quad (2)$$

where  $b_k$  is the partial sill and  $g_k(\mathbf{h})$  is the variogram model of the  $k$ th structure.

A very important property of the linear model of regionalization is that it can be used to decompose the original random function  $Z(x)$  into  $q$  independent random functions (called spatial components) corresponding to spatial scale  $k$ . Decomposition is based on the following model:

$$Z(x) = \sum_{k=1}^q a_k Y_k(x) + m \quad (3)$$

where  $a_k$  are known coefficients and  $Y_k(x)$  are the orthogonal spatial components with spatial covariance  $c_k(\mathbf{h})$ :

$$E\{Z(x)\} = m$$

$$E\{Y_k(x)\} = 0$$

$$\text{Cov} \{Y_k(x), Y_{k'}(x+\mathbf{h})\} = \begin{cases} c_k(\mathbf{h}) & \text{if } k = k' \\ 0 & \text{otherwise} \end{cases}$$

This decomposition is of important practical interest: it makes the distinction of the original random function  $Z(x)$  into  $q$  uncorrelated random functions possible, which represent different scales of variation. Additionally, the estimated spatial components can be mapped using a modified kriging system of equations, which may assist in their interpretation (see also the example in the next Chapter). Mapping spatial components is done with a modified kriging system of equations. Given a set of  $n$  observations, the optimal weights ( $\lambda_\beta$ ) for the estimation of spatial component  $k$  are given by the solution of the following system of equations (ISATIS, 2008):

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta \gamma_{\alpha\beta} + \mu = \gamma_{\alpha 0}^k \\ \sum_{\beta=1}^n \lambda_\beta = 0 \end{cases} \quad a = 1, \dots, n \quad (4)$$

where  $\gamma_{\alpha\beta}$  is the semivariance between data locations  $a$  and  $\beta$ , and  $\gamma_{\alpha 0}^k$  is the semivariance between location  $a$  and the point where an estimation is required. System [4] is identical to a usual kriging system of equations, save the term  $\gamma_{\alpha 0}^k$ , which is the semivariance computed using only the variogram model corresponding to the  $k$ th structure. Another difference is that kriging weights  $\lambda_\beta$  must sum to zero (unlike ordinary kriging where weights sum to unity). This difference is due to the fact that the mean of the random function  $Z(x)$  is considered a part of the spatial component with the largest range [ $\max(k)$ ]. In this case (i.e., when estimating the largest range component) the system [4] should be rewritten in order to account for the varying spatial mean of the function (Isatis, 2008):

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta \gamma_{\alpha\beta} + \mu = \gamma_{\alpha 0}^{\max(k)} \\ \sum_{\beta=1}^n \lambda_\beta = 1 \end{cases} \quad a = 1, \dots, n \quad (5)$$

### 2.1.2 Example: A model of regionalization for zinc in the Duero basin

The sample variogram for Zn concentration in Duero basin soil samples presents two different slopes before reaching the sill at a distance of approximately 120 km (Figure 5). Apparently, the total spatial variability of this variable is structured around two spatial scales at 20 km (local) and at 120 km (regional). The model of regionalization adjusted to the sample variograms of Zn in the Duero Basin is composed of three structures, namely a nugget effect model and two spherical models, with 20 km and 120 km ranges of spatial correlation. The spatial components corresponding to this model may now be used to decompose the original variable into two components by separating the variation observed on a small spatial scale from the one on a larger scale (see the maps in Figure 5).

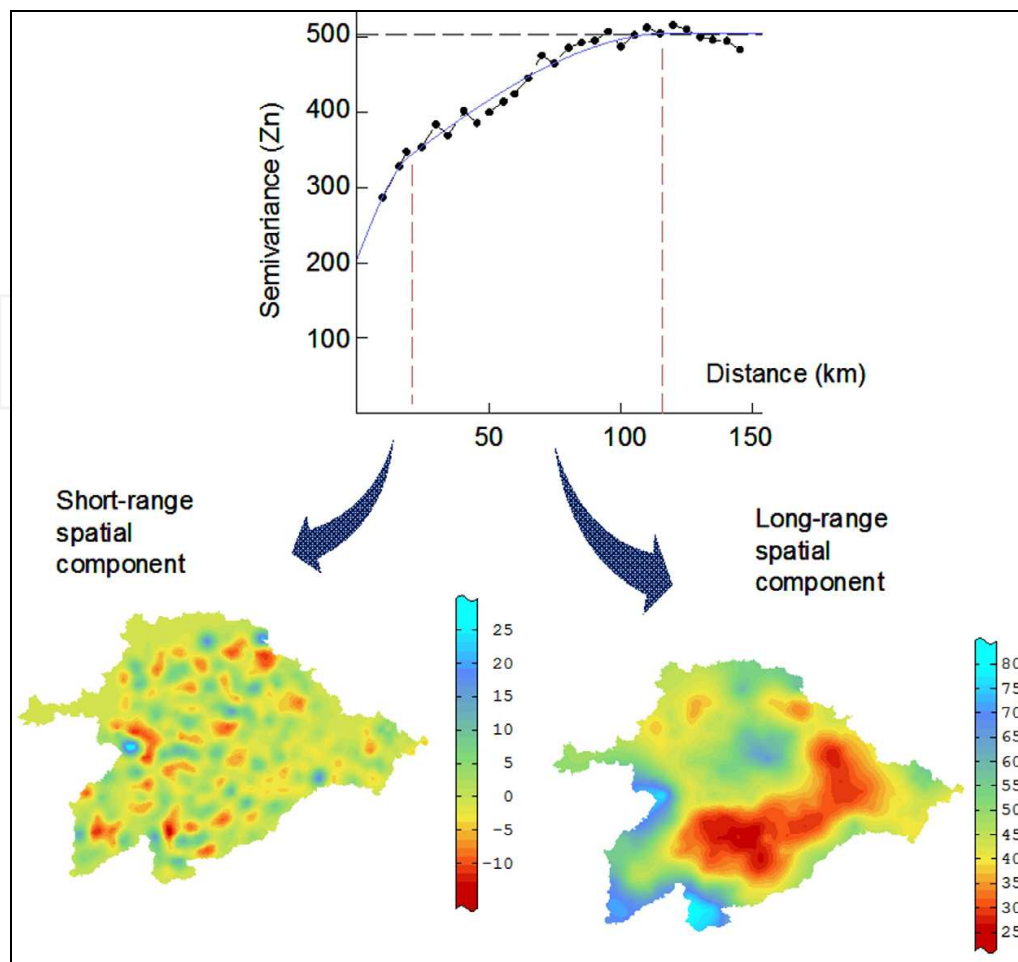


Fig. 5. Mapping two spatial components for Zn concentration in Duero basin soils. The upper panel shows the linear model of regionalization adjusted to the experimental variogram. The lower panels present the estimated spatial components with short- (left panel) and long- (right panel) range variation.

The variation of the component with the longest range (120 km in this case) in the distribution of heavy metals in soils tends to attribute to geological-type factors of a natural origin. Rock decomposition and its input of elements to soil formation (Alloway, 1995) are two main factors that influence on this scale. Hence, large lithological units determine zinc distribution on this scale in the Duero valley. On the other hand, zinc is not a metal related with atmospheric deposition processes over long distances, and is not associated with an industrial activity that might influence the Duero valley on such a large scale. An alternative interpretation of the influence of farming treatments on zinc content in soil is limited to the extension of plots used for crop-growing. The spatial extent of the spatial component in Figure 5 is much greater than the extended land use polygons in the study area.

Unlike the long-range component, its counterpart has a range of spatial correlation of just 20 km, which may be attributable to both antropogenic and natural factors. To a minor spatial extent, zinc in soil would be related with edaphic parameters, such as organic matter, clays, cationic exchange capacity, or presence of oxides of other metals like Fe or Al, which favor its metals retention and whose variation in soil may correspond to the ranges of this scale. On the

other hand, the variability on this small-range scale (20 km) may also reflect antropogenic activities in either agriculture or small industrial plants. With this kind of statistical analysis, it is impossible to distinguish the repercussion or influence of both factors (human input from natural input) beyond assumptions. However, and in relation to some low Zn contents (save the odd exception), it would be possible to rule out that inputs originating from generic industrial activities, spillages or sporadic pollution sources would be of much influence on this scale. Nonetheless, it would be necessary to know the relationship with other edaphic parameters, specifically with those metals that might relate more with farming activities.

## 2.2 Multivariate analysis

The natural extension of the variogram to include two random functions ( $i$  and  $j$ ) is called the cross-variogram between  $Z_i$  and  $Z_j$ , which is estimated as:

$$\hat{\gamma}_{ij}(h) = \frac{1}{2N(h)} \sum_{x_\beta - x_\alpha \approx h} [Z_i(x_\beta) - Z_i(x_\alpha)][Z_j(x_\beta) - Z_j(x_\alpha)] \quad (6)$$

where all the symbols are as in (1) but using the subscript  $i$  or  $j$  to distinguish between the two variables.

Variogram modeling is more complicated in the multivariate case since one needs to model a total of  $p(p+1)/2$  direct and cross-variograms (where  $p$  is the number of variables). Multivariate variogram fitting is usually accomplished with the use of multiple nested structures where each one corresponds to a different scale of variation. The model –called the linear model of coregionalization, LMC– can be written as:

$$\gamma_{ij}(\mathbf{h}) = b_{ij}^{(1)} g^{(1)}(\mathbf{h}) + b_{ij}^{(k)} g^{(k)}(\mathbf{h}) + \dots + b_{ij}^{(q)} g^{(q)}(\mathbf{h}) \quad \forall i, j \quad (7)$$

where  $\gamma_{ij}(\mathbf{h})$  is the variogram model for variables  $i$  and  $j$  (for  $i=j$  the auto-variogram is obtained),  $b_{ij}^{(k)}$  is the partial sill for the  $ij$ th variogram for structure  $k$ , while  $g^{(k)}(\mathbf{h})$  represents the type of variogram model (i.e., exponential, spherical, etc.) for structure  $k$ . The first structure  $g^{(1)}$  represents the nugget effect model.

The flexibility of a specific LMC (that is, its ability to model a set of experimental variograms) is based on the total number of basic structures ( $q$ ) and their corresponding range of spatial correlation, as well as the partial sill  $b_{ij}^{(k)}$  for each variogram model and structure. Partial sills  $b_{ij}^{(k)}$  may vary across variograms (under some restrictions; see Wackernagel (1998)), but the range of spatial correlation of each structure  $g^{(k)}(\mathbf{h})$  should be the same for the set of  $p(p+1)/2$  variogram models. If we arrange the partial sills  $b_{ij}^{(k)}$  of the LMC in a matrix form, we obtain the so-called co-regionalization matrices. The coregionalization matrix for structure  $k$  is a positive semi-definite symmetric ( $p \times p$ ) matrix with diagonal and off-diagonal elements of the partial sill of the auto- and cross-variogram models, respectively, obtained from the LMC:

$$\mathbf{B}_k = \begin{bmatrix} b_{11}^{(k)} & \dots & b_{1p}^{(k)} \\ \dots & \dots & \dots \\ b_{p1}^{(k)} & \dots & b_{pp}^{(k)} \end{bmatrix} \quad (8)$$

Note that the LMC is a permissible model, but only when all the co-regionalization matrices  $\mathbf{B}_k$  are positive definite. This constraint makes multivariate variogram fitting a difficult task. Should only two variables be implicated ( $p=2$ ), the linear model of coregionalization can be fitted manually. However for  $p>2$ , a weighted least squares approximation is needed which offers the best LMC fit under the constraint of positive semi-definiteness of  $\mathbf{B}_k$  (Goulard and Voltz 1992). However, the use of the LMC facilitates multivariate variogram modeling since it reduces the problem of fitting a total of  $p(p-1)/2$  variogram models to their experimental counterparts to the problem of deciding the total number ( $q$ ) of structures to use, as well as the type (i.e., spherical, exponential, etc.) of each structure. This is a central decision in the MFK framework. There are no general rules to guide this decision; however, some recommendations may prove useful (see also Goulard and Voltz (1992)):

- Experimental variograms are the basis for deciding the number and range of elementary structures to be used for modeling. One usually tries to find subsets of variables that have similar variogram characteristics (equal range of spatial correlation).
- In practice, it is difficult to distinguish more than three structures (a nugget plus two other structures) in a set of experimental variograms. Practical experience has shown that three basic structures are sufficient for modeling a large number of variables.
- Usually some of the original variables in the dataset should share the same structures.
- Specifically for heavy metal distribution in soils, geological maps can prove most helpful for deciding the number and ranges of the spatial correlation of individual model structures. The spatial distribution of heavy metals in soil has been observed to be in close relationship with the basic geological features –and their spatial extent– of the study area.

### 2.2.1 Example: Fitting the linear model of coregionalization in the Ebro river basin

Let us now describe the procedure for fitting a linear model of coregionalization for the spatial distribution of heavy metals in the Ebro basin (Rodríguez et al., 2008). By looking at the direct experimental variograms (Figure 6), we note a change in their slope at a distance of approximately 20 km, which is more prominent for some metals (Hg and Cd) than for others (Zn). This common feature in all the sample variogram indicates that the linear model of coregionalization should include a short-range structure. Note also that all the sample variograms seem to reach the sill value at a distance of approximately 200 km, a fact which also makes necessary the inclusion of a long-range component in the model of coregionalization to be built. Finally, we note that the ascension in the semivariance observed between the shortest range (20 km) and the longest one (200 km) is not linear; instead we observe a stabilization at distances of approximately 100 km.

The features observed in the sample variograms act as a basis for deciding that the linear model of coregionalization to be postulated –and fitted afterward to the sample variograms–

should be composed of at least three structures with ranges of approximately 20 km, 100 km and 200 km (plus a nugget effect model which is present in all the sample variograms). Cross variograms - not presented here - were also taken into account when building the model. However, the postulated model should be chiefly designed to fit the direct rather than the cross-variograms as precisely as possible.

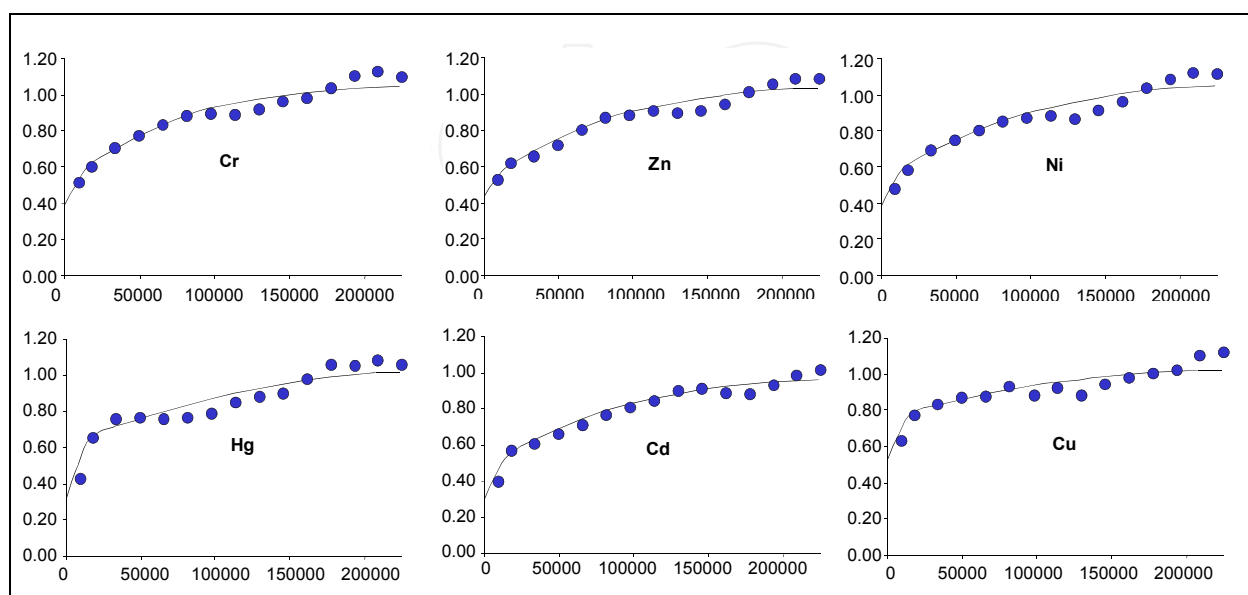


Fig. 6. Six of the seven variogram models of the linear model of coregionalization for the concentration of heavy metals in Ebro basin soils. Horizontal axes units are provided in meters, while vertical axes represent the standardized -to unit sill- semivariance of the corresponding element.

In Figure 6, we graphically present the linear model of coregionalization fit after an iterative procedure included in the Isatis software (Isatis, 2008). The fit is not actually perfect, especially for some elements such as Hg. Yet when considering a multiple variogram model, a compromise between model simplicity and accuracy of fit should be reached.

The number and range of the structures adopted in this model are coherent from a practical viewpoint. Note that within the study area, we have reported the presence of several industrial activities that may potentially enrich the basin soil through airborne pollution and subsequent deposition. We expect that this activity may have altered the spatial distribution at intermediate distances from the point source. Mercury emitted from industrial plants, for instance, has been reported to travel tens of thousands of kilometers before being deposited in soil. On the other hand, large geological features within the study area make us think that natural geological variation could be responsible for the spatial distribution of some heavy metals on a large spatial scale (that with the longest range). Finally, differences in land use and short-range scale natural processes affecting pedogenesis are believed to act on short spatial scales, and are presumably responsible for the short-range scale effects in the distribution of soil heavy metals.

In a very similar way, we built the linear model of coregionalization for heavy metals concentration in Duero basin soils. The sample variograms features in this basin are similar: variograms present a short-range structure at a distance of 20 km and a medium-range

structure at 100 km – 120 km. A substantial difference, however, between the two basins is noted in the absence of long-range variation in the Duero basin. Therefore, the long-range structure (220 km) found in the Ebro region is absent in this model (see also Figure 7).

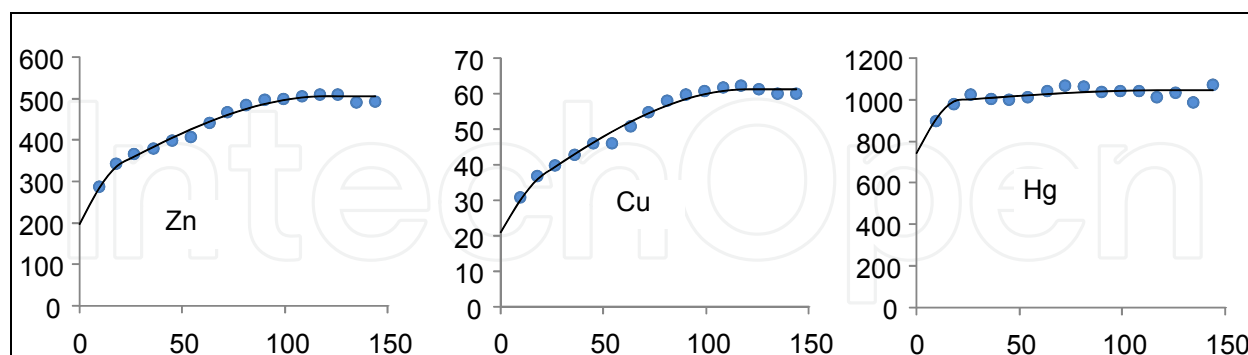


Fig. 7. Three of the seven variogram models of the linear model of coregionalization for the heavy metals concentration in Duero basin soils. Horizontal axes units are provided in kilometers, while vertical axes represent the semivariance of the corresponding element.

### 2.2.2 Extracting spatial components

Analogously to the univariate case, the LMC permits the decomposition of the original random functions  $Z_i(x)$  into a linear combination of the  $q$  mutually uncorrelated random functions  $Y_{lk}$ , called regionalized factors (RFs):

$$Z_i(x) = m_i(x) + \sum_{k=1}^q \sum_{l=1}^p a_{il}^k Y_{lk}(x) \quad i = 1, \dots, p$$

or in matrix form:

$$\mathbf{Z} = \mathbf{m} + \sum_{k=1}^q \mathbf{A}_k \mathbf{Y}_k \quad (9)$$

where  $m_i(x)$  is the varying mean of the function and  $\mathbf{A}_k$  is a matrix of unknown coefficients. In fact, only the coregionalization matrix for the  $k$ th structure:

$$\mathbf{B}_k = \mathbf{A}_k \mathbf{A}_k' \quad (10)$$

can be estimated. Matrix  $\mathbf{A}_k$  is not uniquely defined since there is an infinite number of  $\mathbf{A}_k$  which satisfies Equation (10). However, a principal components analysis (PCA) can provide a natural determination of matrix  $\mathbf{A}_k$  (Chilés and Delfiner 1999). More specifically, a PCA is applied to each coregionalization matrix  $\mathbf{B}_k$  separately, to provide a set of  $p$  eigenvalues and their corresponding eigenvectors ( $\mathbf{u}_l$ ):

$$\mathbf{B}_k = \mathbf{Q}_k \mathbf{\Lambda}_k \mathbf{Q}_k' \quad (11)$$

where  $\mathbf{Q}_k$  is the orthogonal matrix of eigenvectors and  $\mathbf{\Lambda}_k$  is the diagonal matrix of eigenvalues for spatial scale  $k$ .

Matrix  $\mathbf{A}_k$  can then be estimated since:

$$\mathbf{A}_k = \mathbf{Q}_k \mathbf{\Lambda}_k^{1/2} \quad (12)$$

Using decomposition (9) and the estimation of matrix  $\mathbf{A}_k$  from (12), one can decompose the  $p$  original variables into a set of uncorrelated RFs. Note that these factors have some remarkable properties:

- They are linear combinations of the  $p$  original variables.
- The first factors, corresponding to the largest eigenvalues, account for most of the variability observed in the dataset. Therefore, it is possible to reduce the dimensionality of the data and to visualize the phenomenon.
- It is possible that the resulting RFs represent some common mechanisms underlying the spatial distribution of the original variables. Hopefully, RFs can have meaningful interpretations (analogically to a classical PCA).
- RFs remain uncorrelated at any separation distance  $\mathbf{h}$  and not just at the same location (Chilés and Delfiner 1999). A classical PCA provides uncorrelated factors, but only for the separation distance  $\mathbf{h}=0$ .
- Perhaps the most remarkable property is that RFs can be constructed for each spatial scale separately. Therefore, by incorporating the spatial correlations revealed by the variograms, they can reveal mechanisms that act on different scales and that control the spatial distribution of the studied attributes.

### 2.2.3 Interpreting the regionalized factors: The circle of correlation

Interpretation of an RF is not always feasible given that RFs are determined using statistical and not ecological/physical criteria (maximize variance under the constraint of orthogonality). Nevertheless, the interpretation of RFs is of crucial importance for the analysis since it will be the basic tool for determining the exact physical mechanisms acting on different spatial scales and for controlling the spatial distribution of the studied variables.

The interpretation of the meaning of an RF can be assisted by all the well-known tools of a classical PCA, such as “scree” plots, loadings, and especially by the correlation between the regionalized factors and the regionalized variables, which is computed by:

$$\rho_{il} = q_{il} \sqrt{\frac{\lambda_l}{\sigma_i^2}} \quad (13)$$

where  $q_{il}$  is the loading of the  $i$ th variable for the  $l$ th principal component,  $\lambda_l$  is the variance of the  $l$ th RF and  $\sigma_i^2$  the variance of the  $i$ th regionalized variable.

The pair of correlations of a regionalized variable with two RFs (usually the first two, which account for most of the variability) is plotted on a graph, called the circle of correlation (Saporta, 1990). When a variable is located near one of the two axes of the graph and away from the origin, it is well correlated with that specific RF and much of its variance is explained by the RF.



### 2.2.4 Example: Multiscale correlations among heavy metals in the Ebro river basin

In this Chapter, we use the linear model of corregionalization to decompose the initial multivariate set of heavy metal concentrations. Then we derive a new set of composite regionalized factors for several scales of variation. Finally, we employ a factorial kriging analysis results in an attempt to decompose the original dataset into a few regionalized factors.

For the Ebro river, three spherical models with ranges of 20 km, 100 km and 220 km were used. The three circles of correlation shown in Figure 8 represent heavy metal associations for the three scales of variation used in the linear model of corregionalization for soil samples. The association of different metals substantially differs if we compare the three circles of correlation. This indicates that, indeed, the correlation among soil elements depends on the spatial scale considered and that, conclusively, multiscale correlation is present in the study area. Grouping metals in the circle of each scale tends to reflect the influence phenomena that are common for all grouped elements. Changes made to the grouping of these metals when amending the observation scale also implies a change in the dominant factor which influences these elements. Note that in the absence of multiscale correlation, we expect the elemental associations to remain unchanged when moving from one spatial scale to another.

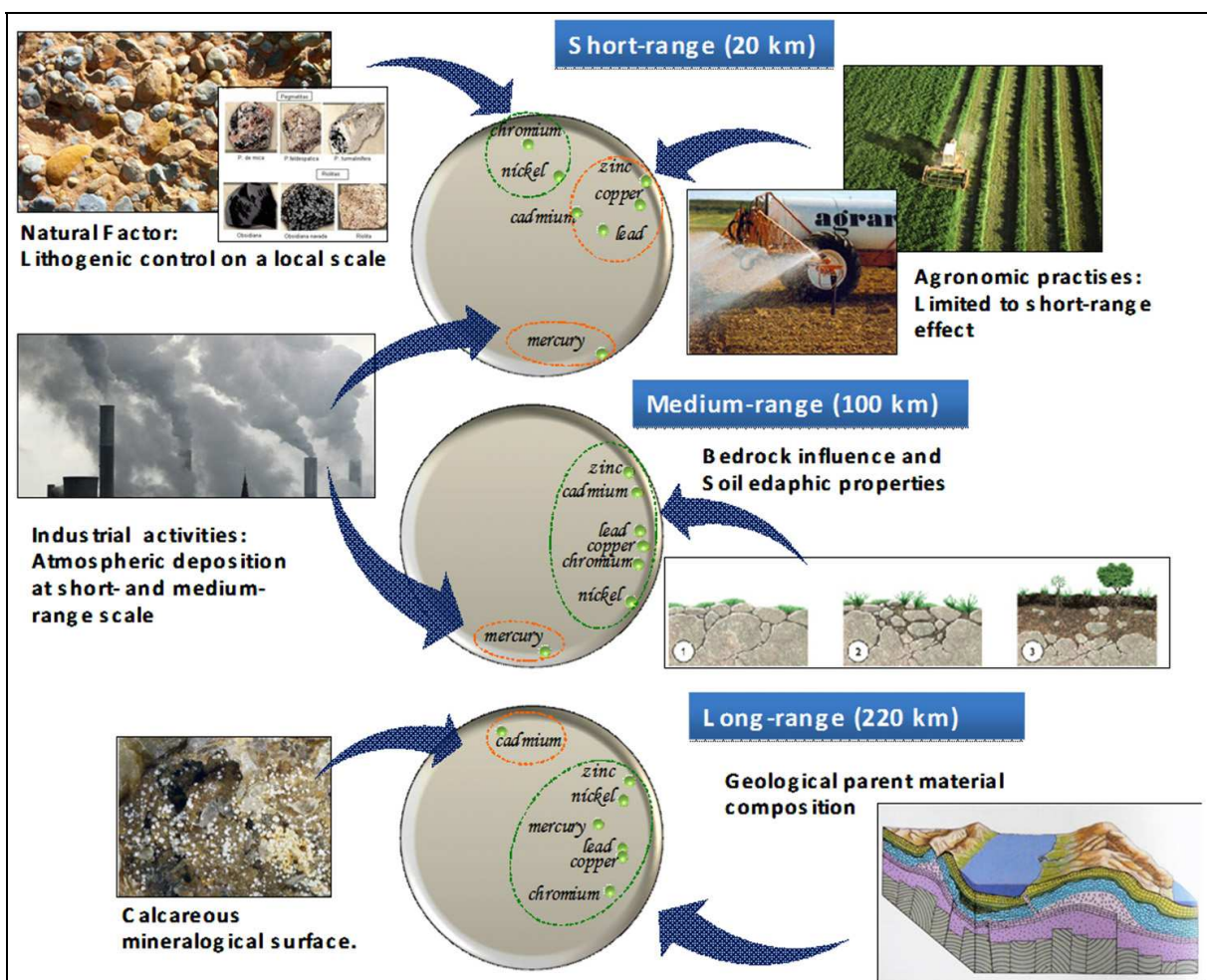


Fig. 8. Correlation circles for the concentration of seven heavy metals in the Ebro river basin. Correlation circles correspond to the three structures used to build the linear model of corregionalization.

The association between Hg and other elements is the most representative example of a correlation dependent on scale. Note that Hg seems to be an isolated element in the correlation circles for the short- and medium-range correlation circles, but seems to be associated with other elements when looking at the long-range correlation circles. Obviously, there should be a factor that acts on the smallest scale of variation, but not on the largest one, which makes mercury distribution different from that of the other elements. In general, mercury accumulations in soils are associated with atmospheric deposition (Engle et al., 2005). Anthropogenic emission of Hg represents about 60–80% of global Hg emissions. Mercury is an extremely volatile metal which can be transported over long distances (Navarro et al., 1993). In the Ebro valley, soil enrichment by atmospheric mercury covers small (20km) to medium (100km) scales. Not surprisingly, soil enrichment in mercury is not observed on the largest (among those studied) spatial scales, so man-made alteration of soil Hg has not become evident –at least not yet- on this scale.

Cd is another interesting element which shows correlations with other elements that change across scales. In the largest scale of variation (220 km), we may denote Cd as an isolated element in the circle of correlation which is, however, not observable on smaller spatial scales. Unlike Hg, however, this change can be attributed to natural factors. More specifically, Cd is the only element that tends to accumulate in calcareous soils (Boluda et al., 1988). Cadmium is adsorbed specifically by crystalline and amorphous oxides of Al, Fe and Mn. Metallic (copper, lead and zinc); alkaline earth cations (calcium and magnesium) particularly reduce Cd adsorption by competing for available specific adsorption and cation-exchange sites (Martin and Kaplan, 1998). Therefore, the isolated occurrence of Cd in the long-range correlation circle is due to the spatial distribution of the calcareous mineralogical surface generating major Cd accumulation.

On the local spatial scale (20km), association of Zn, Cu, Pb and Cd results concentrations after fertilization of arable soils, or pesticides and fungicides use related to crop protection. The association of four elements is caused by common agricultural practices, which may also prove evident on a small scale. Copper and zinc (Mantovi et al., 2003) increase through such practices especially copper which has been used as a pesticide form (copper sulfate) in viticulture. Zinc and cadmium concentrations increase through use of fertilizers. It is estimated that phosphated fertilizers make up more than 50% of total cadmium input in soils (de Meeûs et al., 2002). On the other hand, natural phenomena act in a short range, and variability of Cr and Ni is associated with the mineralogical structure of the study area (basic and ultramafic rock). Generally, anthropic inputs of Cr and Ni in fertilizers, limestone and manure are lower than the concentrations already present in soil (Facchinelli et al., 2001).

### **2.2.5 Example 2: Multiscale correlations among soil heavy metals in the Duero river basin**

The linear model of co-regionalization in the Duero basin is developed from two structures, with scale ranges between 20 km and 120 km. Unlike the Ebro valley, no higher scale of variation is noted, which is around 200 km, despite having a similar surface. The Duero valley can be considered more homogeneous in terms of all the aspects that can influence the distribution of heavy metals in soils, and reducing spatial variation structures may well reflect this. Moreover, the Duero depression is shaped as a basin with tertiary and

quaternary sediments of a continental origin. However, Paleogene materials of variable extensions outcrop from among the tertiary sediments, although they mainly limit the depression on the basin's edges. Variability from a geological viewpoint is inferior to variability in the Ebro. Yet from the farming perspective, crops do not present much diversity as dominant crops are cereals and vineyards. As mentioned in Section 1.4.4, the Duero valley is one of the main cereal-growing areas in Spain.

On the short-range scale (20km), only two groups of metals are seen, which are expressed as mercury isolation and a large series with the remaining elements (Figure 9). The main difference found with the Ebro valley lies in chrome and nickel grouping with metals like copper or zinc, which relate to farming practices. Although natural and inorganic fertilizers contain small amounts of both chrome and nickel, they do not significantly increase the soil natural content. The geological influence of nickel and chrome content has been clearly demonstrated (Alloway, 1995). The highest nickel concentrations are found in ultrabasic igneous rocks (peridotites, dunites and pyroxenites), followed by basic rocks (gabbro and basalt). Acid igneous rocks present lower chrome or nickel contents. Sedimentary rocks are especially poor in Cr or Ni. On the other hand, the edaphic parameters that influence heavy metals content in soil (texture, organic matter, etc.) evidence their influence on this scale. Soils with a thick and sandy texture contain less Cr or Ni than clayey soils, which is also a common process for the remaining metals (Cu, Pb, Cd and Zn). As expected, low heavy metal contents are also associated with low organic matter contents.

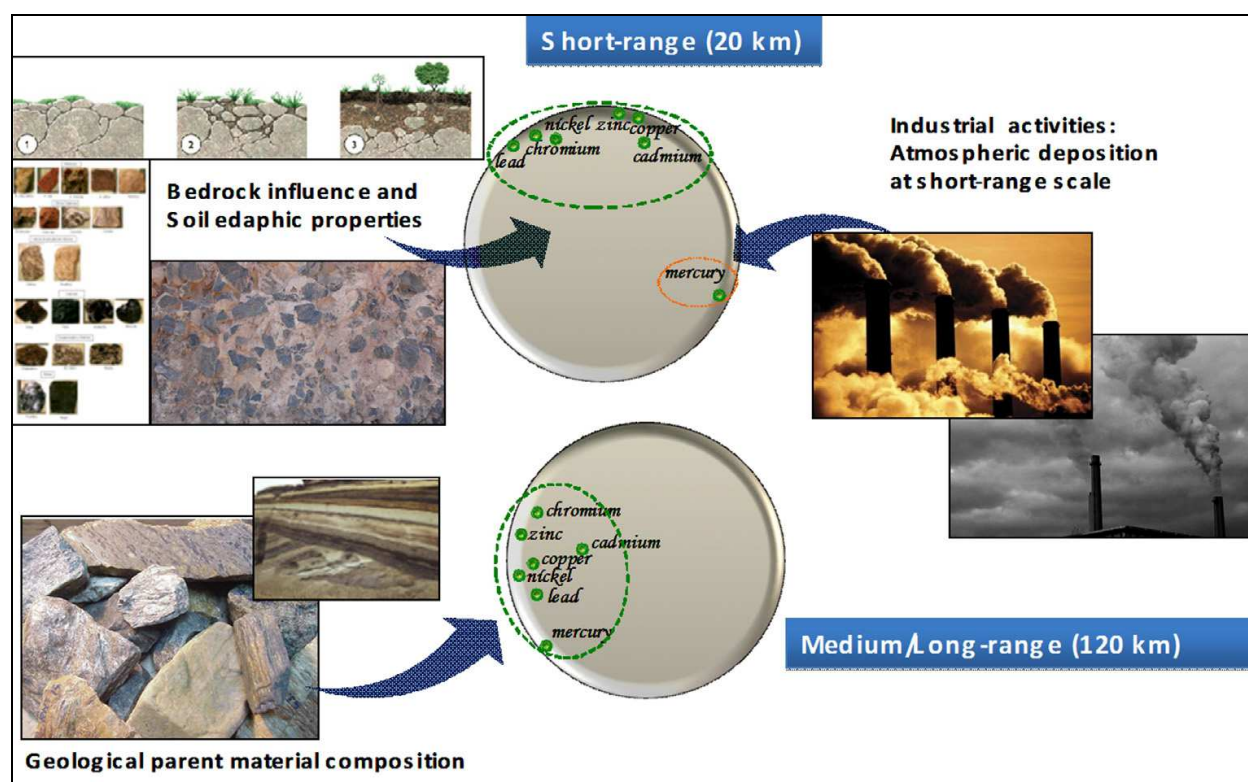


Fig. 9. Correlation circles for the concentration of seven heavy metals in the Duero basin. Correlation circles correspond to the two structures used to build the linear model of corregionalization.

On the medium-/long-range scales (120km), all heavy metals are grouped in the circle of correlation (Figure 9), which is the result of a common source of variability. On this scale, only large lithologies can influence the distribution of heavy metals in soils which, for the River Duero basin, derive mainly from the sedimentary materials deposited in its interior. However, the coordinated distribution of all metals, including Hg, on this scale evidences an edaphogenic process.

In conclusion, human activity is only shown on a small scale (20km) and is motivated by mercury inputs. As mentioned earlier, atmospheric deposition is the main way that mercury enters soil. Besides, on this small scale, it is probably a case of mercury being associated with ash and soot particles, which are deposited near pollution sources, such as those originating from coal power stations, or from heating systems to a lesser extent. Basically, no influence of farming treatments is observed. This aspect is linked to some relatively low levels of metals in soil, such as Cu, Cd or Zn. The dominant factor in the content and distribution of heavy metals in the Duero is natural, and is influenced by soil's physico-chemical properties in a short range and, to a greater extent, by geological parent material composition.

### 2.2.6 Spatial estimation of regionalized factors

The final step in factorial kriging analysis consists in mapping composite regionalized factors. Estimation is done with a modified cokriging step technique. Note that ordinary cokriging provides a spatial estimation of the primary variable using data for the primary variable and for one or more secondary variables. Typically, the primary variable is sampled over a limited number of points, while secondary data are more densely sampled (Wackernagel 1998). Estimation of regionalized factors is done by means of a modified cokriging system of equations, where measurements of the primary data are not available, and the cross covariance between the regionalized factor and the regionalized variables cannot be inferred directly. However, the spatial estimation of regionalized factors is possible thanks to the model of coregionalization's properties. More specifically, to account for decomposition [8], the cross covariance between the random variable  $Z_i(x)$  and the  $l$ th RF of the  $k$ th spatial structure ( $Y_{lk}$ ) is determined as (Goovaerts 1997):

$$\text{Cov}\{Z_i, Y_{lk}\} = a_{il}^k c_{a_i,0}^k \quad (14)$$

where  $c_{a_i,0}^k$  is the covariance (for the  $k$ th structure) between point  $a_i$  (where the  $i$ th variable has been measured) and the location where the prediction is required, while  $a_{il}^k$  is the element of the  $i$ th row of the  $l$ th column of matrix  $\mathbf{A}_k$  determined by (12). The optimal cokriging weights  $\lambda_{\beta,l}^k$  assigned to each data location for each regionalized factor are provided by the solution of the following cokriging system of equations:

$$\begin{cases} \sum_{j=1}^p \sum_{\beta_j=1}^{n_j} \lambda_{\beta_j,l}^k C_{ij}(\mathbf{u}_{\alpha_i} - \mathbf{u}_{\beta_j}) + \mu_{il}^k = a_{il}^k c^k(\mathbf{u}_{\alpha_i} - \mathbf{u}) & \alpha_i = 1, \dots, n_i \quad i = 1, \dots, p \\ \sum_{\beta_i=1}^{n_i} \lambda_{\beta_i,l}^k = 0 & i = 1, \dots, p \end{cases} \quad (15)$$

where  $n_i$  and  $n_j$  are the number of sample locations for the original variable  $i$  and  $j$ , respectively, while  $\sigma_{ij}$  is the auto- (or cross-) covariance for variables  $i$  and  $j$  between locations. Finally,  $\lambda_l$  is the Lagrange multiplier for the  $i$ th variable and the  $l$ th regionalized factor.

The cokriging system in (15) differs from the classical cokriging system as far as the way to compute the cross covariance between primary and secondary data and in the unbiasedness constraints is concerned. More specifically, since the regionalized factors have (by definition) a zero mean, the cokriging weights have to sum to zero for the system to be overall unbiased.

### 2.2.7 Example: Spatial estimation of regionalized factors in the Duero

The first regionalized factor of the long-range scale of variation presents an area of high positive values in the southern part of the Duero basin. This area is characterized by a higher metal concentration for all the analyzed elements. According to the results presented in Section 2.2.3, the basin's great lithological features are responsible for enriching this area with heavy metals. Conclusively, the variability depicted in this map, representing a juxtaposition between metal-rich areas and poorer ones, is caused by the chemical composition of the geological substrate; human sources of heavy metals are not obvious in this map.

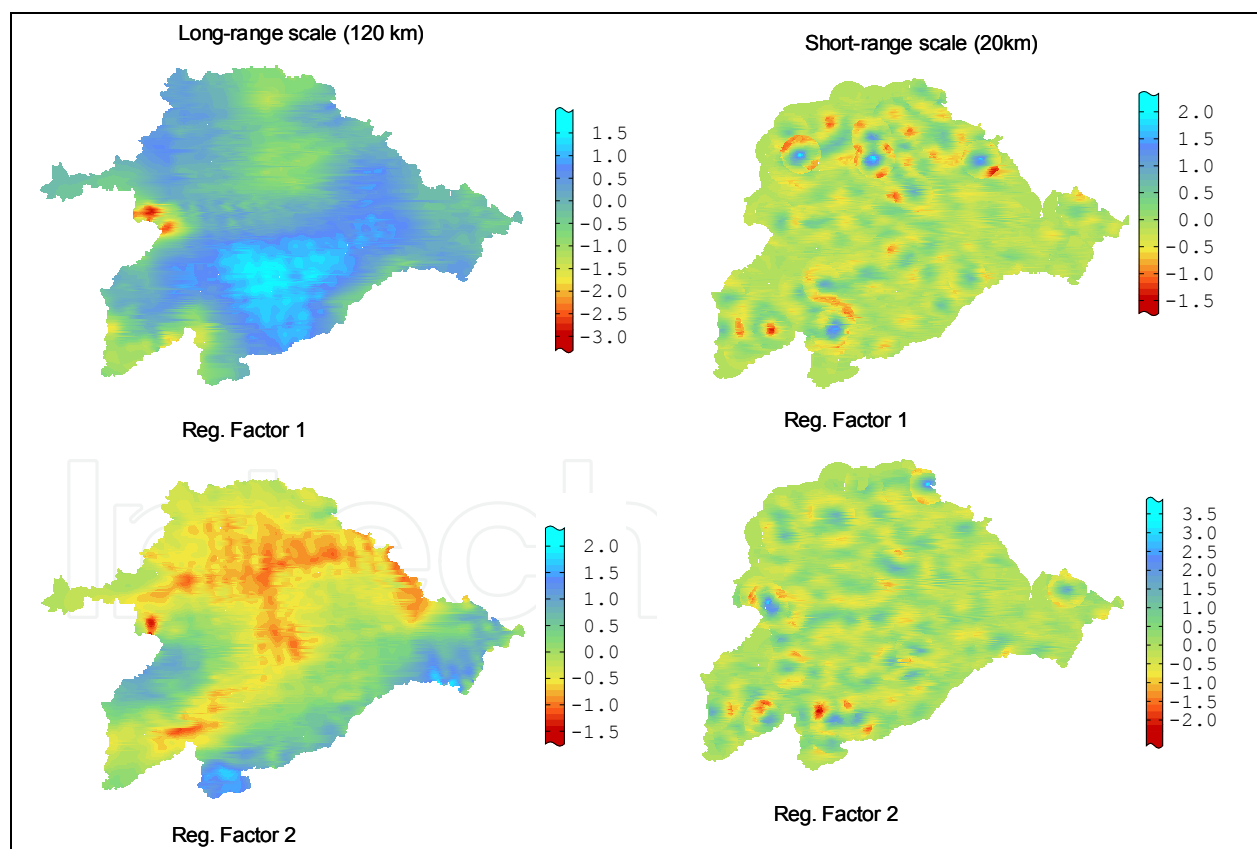


Fig. 10. Spatial estimation of regionalized factors for the Duero case study. Note that several grid points in the short-range spatial components cannot be estimated due to the limited number of data points. Thus, the study areas seem smaller in the maps of the short-range structure.

In contrast, variability in the first regionalized factor of the 20 km range scale is, according to the previous results, due to man-made enrichment of soil with Hg. Therefore, areas with higher regionalized factor values for this map depict those areas where mercury enrichment by human activities has been observed. The second regionalized factor for the same scale of variation depicts small-scale variability in all elements –except the previously reported Hg– which is due to the natural edaphic properties of soil (texture, soil organic matter, etc.) observed on small spatial scales..

### 3. Conclusion

Several human activities, such as agriculture, mining or industry, have dramatically increased the concentration of some heavy metals in soil causing, in some cases, severe soil pollution. Source identification of heavy metals can prove to be a great step forward in the prevention of soil pollution and rationalization of soil management practices. However, when multiple sources of heavy metals contribute to total concentration and are mixed with natural metals in soil, then the task of identifying and assigning pollutant sources remains an unresolved problem.

Fortunately, some human heavy metal sources enrich soil on different spatial scales. For instance, the impact of fertilization of crops in arable soil is on a rather local range. This is the case of Zn, Cu, Pb and Cd on the local scale of variation in the Ebro basin. Conversely, contaminant sources, such as industrial plants, disperse large quantities of heavy metals over long dispersal distances. The most obvious case was detected in both the Ebro and Duero basins for the spatial distribution of Hg. Finally lithological features, which also control the –natural– spatial distribution of heavy metals, may have a much broader range of spatial distribution than human enrichment factors. This was the case in the Ebro basin where Cd distribution on the longest-range spatial scale was found to be controlled by lithology.

A factorial kriging analysis was presented as a tool to provide some evidence about the source identification of soil heavy metals. From a statistical viewpoint, the method relies heavily on a scale-dependent decomposition of the original variables (the heavy metal concentration) in a new set of composite functions (regionalized factors). These new unobservable functions may be interpreted using the correlation circle and spatial distribution maps. Regionalized factors have the remarkable property of being scale-specific. Regionalized factors are constructed based only on the scale-specific correlations among the original variables. Thus they can be used to filter out variability on unwanted spatial scales and to reveal scale-specific common sources of heavy metals in soil.

### 4. Acknowledgment

We appreciate the financial assistance provided by the Spanish Ministry of Innovation through project JC2010-0109. We are also grateful to Ministerio de Ciencia e Innovacion, proyect: CGL2009-14686-C02-02 and to CAM project: P2009/AMB-1648 CARESOIL.

### 5. References

Alloway, B.J. (1995). Heavy metals in soils. Chapman & Hall, (Ed.). Glasgow, UK.

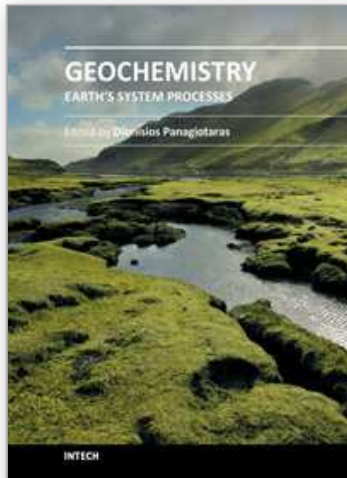
- Alloway, B.J. & Jackson, A.P. (1991). The behaviour of heavy metals in sewage sludge-amended soils. *The Science of the Total Environment*, Vol. 100, pp. 151-176.
- Boluda, R.; Andreu, V.; Pons, V. & Sánchez, J. (1988). Contenido de metales pesados (Cd, Co, Cr, Cu, Ni, Pb y Zn) en suelos de la comarca La Plana de Requena-Utiel (Valencia). *Anales de Edafología y Agrobiología*, Vol. 47, No. 11-12, pp. 1485-1502.
- Brunet, M. F. (1986). The influence of evolution of the Pyrenees on adjacent basins. *Tectonophysics*, Vol. 129, pp. 343-354.
- Castrignanò, A.; Giugliarini L.; Risaliti R. & Martinelli, N. (2000). Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. *Geoderma*, Vol. 97, No 1-2, pp. 39-60
- Chilés, J.P. & Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. John Wiley & Sons (Ed.), New York, 695 pp.
- Chojnacka, K.; Chojnacki, A.; Górecka, H. & Górecki, H. (2005). Bioavailability of heavy metals from polluted soils to plants. *The Science of the Total Environment*, Vol. 337, pp. 175-182.
- Colgan, A.; Hankard, P.K.; Spurgeon, D.J.; Svendsen, C.; Wadsworth, R.A. & Weeks, J.M. (2003). Closing the loop: A spatial analysis to link observed environmental damage to predicted heavy metal emissions. *Environmental Toxicology and Chemistry*, Vol. 22, pp. 970-976.
- De Meeûs, C.; Eduljee, G. & Hutton, M. (2002). Assessment and management of risks arising from exposure to cadmium in fertilizers. *The Science of the Total Environment*, Vol. 291, pp 167-187.
- De Temmerman, L.; Vanongeval, L.; Boon, W.; Hoenig, M. & Geypens, M. (2003). Heavy metal content of arable soils in northern Belgium. *Water, Air and Soil Pollution*, Vol. 148, pp. 61-76.
- Dudka, S.; Piotrowska, M. & Chlopecka, A. (1994). Effect of elevated concentrations of Cd and Zn in soil on spring wheat yield and the metal contents of the plants. *Water, Air, and Soil Pollution*, Vol. 76, pp. 333-341.
- Engle, M.A.; Gustin, M.S.; Lindberg, A.W. & Ariya, P.A. (2005). The influence of ozone on atmospheric emissions of gaseous elemental mercury and relative gaseous mercury from substrates. *Atmospheric Environment*, Vol. 39, pp. 7506-7517.
- Errecalde, M.F.; Boluda, R.; Lagarda, M.J. & Farre, R. (1991). Indices de contaminación por metales pesados en suelos de cultivo intensivo: aplicación en la comarca de L'Horca (Valencia). *Suelo y Planta*, Vol.1, pp. 483-494.
- Facchinelli, A.; Sacchi, E. & Mallen, L. (2001). Multivariate statistical and GIS-based approach to identify heavy metal sources in soil. *Environmental Pollution*, Vol. 114, pp. 313-324.
- Gómez, J.J.; Lillo, J. & Sahún, B. (2006). Naturally occurring arsenic in groundwater and identification of the geochemical sources in the Duero Cenozoic Basin, Spain. *Environmental Geology*, Vol. 50, pp. 1151-1170.
- Goovaerts, P. (1992). Factorial kriging analysis: A useful tool for exploring the structure of multivariate spatial information. *Journal of Soil Science*, Vol. 43, pp. 597-619.
- Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Applied geostatistics series. Oxford University Press (Ed.), New York.
- Goulard, M. & Voltz, M. (1992). Linear coregionalization model: tools for estimation and choice of cross variogram matrix. *Mathematical Geology*, Vol. 24, No 3, pp. 269-286.

- Gzyl, J. (1999). Soil protection in Central and Eastern Europe. *Journal of Geochemical Exploration*, Vol. 66, pp. 333-337.
- Houba, V.J.G.; Van Der Lee, J.J. & Novozamsky, I. (1995). Soil and plant analysis, a series of syllabi. Part 5B: In: *Soil analysis procedures, other procedures*. Dept. of Soil Science and Plant Nutrition, Agricultural Univ (Ed.), Wageningen.
- Hutton, M. (1982). Cadmium in European Community, In: *MARC Report N° 2*, MARC (Ed.), London.
- Isatis (2008). Isatis software manual. Geovariances & Ecole des Mines de Paris, 579 pp.
- ISO 11466. 1995. Soil quality: Extraction of trace elements soluble in aqua regia, Genova, ISO.
- Kashem, M.A. & Singh, B.R. (2001). Metal availability in contaminated soils: Effects of flooding and organic matter on changes in Eh, pH and solubility of Cd, Ni and Zn. *Nutrient Cycling in Agroecosystems*, Vol. 61, pp. 247-255.
- Mantovi, P. ; Bonazzi, G. ; Maestri, E. & Marmiroli, N. (2003). Accumulation of copper and zinc from liquid manure in agricultural soils and crop plants. *Plant and Soil*, Vol. 250, pp. 249-257.
- Matheron, G. (1982). Pour une analyse krigéante de données régionalisées, Centre de Géostatistique, Fontainebleau.
- Martin, H.W. & Kaplan, D.I. (1998). Temporal changes in cadmium, thallium, and vanadium mobility in soil and phytoavailability under field conditions. *Water, air and soil pollution*, Vol. 101, pp. 399-410.
- Navarro, M.; López, H.; Sánchez, M. & López, M. (1993). The effect of industrial pollution on mercury levels in water, soil, and sludge in the coastal area of Motril, Southeast Spain. *Archives of Environmental Contamination and Toxicology*, Vol. 24, pp. 11-15.
- Nicholson, F.A.; Smith, S.R.; Alloway, B.J.; Carlton-Smith, C. & Chambers, B.J. (2003). An inventory of heavy metal input to agricultural soil in England and Wales. *The Science of the Total Environment*, Vol. 311, pp. 205-219.
- Nriagu, J.O. (1990). Global metal pollution. Poisoning the biosphere?. *Environment*, Vol. 32, pp. 28-33.
- Otte, M.L.; Haarsma, M.S.; Broekman, R.A. & Rozema, J. (1993). Relation between heavy metal concentrations and salt marsh plants and soil. *Environmental Pollution*, Vol. 82, pp. 13-22.
- Riba, O.; Reguant, S. & Villena, J. (1983). Ensayo de síntesis estratigráfica y evolutiva de la cuenca terciaria del Ebro. In: *Geología de España*, pp. 41-50, IGME Jubilar J. M. Ríos.(Ed.), Madrid, Spain.
- Ross, S.M. (1994). Toxic metals in Soil-Plant Systems. John Wiley and Sons Ltd. (ed SM Ross). Chichester.
- Rodríguez Martín, J.A.; López Arias, M. & Grau Corbí, J.M. (2006). Heavy metal contents in agricultural topsoils in the Ebro basin (Spain). Application of multivariate geostatistical methods to study spatial variations. *Environmental Pollution*, Vol. 144, pp. 1001-1012.
- Rodríguez Martín, J.A.; López Arias, M. & Grau Corbí, J.M. (2009). Metales pesados, materia orgánica y otros parámetros de los suelos agrícolas y de pastos de España. MARM-INIA (Ed.). ISBN 978-84-491-0980-5. Madrid. Spain.



- Rodríguez, J.A.; Nanos, N.; Grau, J.M.; Gil, L. & López-Arias, M. (2008). Multiscale analysis of heavy metal contents in Spanish agricultural topsoils. *Chemosphere*, Vol. 70, pp. 1085–1096.
- Roure, F.; Chokroune, P. & Deramond, J. (1989). ECORS deep seismic data and balanced cross-sections, geometric constraints to trace the evolution of the Pyrenees. *Tectonics*, No 8, pp. 41-50.
- Sanchidrian, J.R. & Mariño, M., (1980). Estudio de la contaminación de suelos y plantas por metales pesados en los entornos de las autopistas que confluyen en Madrid. II contaminación de suelos. *Anales de Edafología y Agrobiología*, Vol. 39, pp. 2101-2115.
- Saporta, A. (1990). Probabilités, analyse des données et statistique. Technip, Paris.
- Simón, J. L. & Soriano, M. A. (1986). Diapiric deformations in the Quaternary deposits of the central Ebro basin, Spain. *Geological Magazine*, Vol. 123, pp. 45-57.
- Simon-Gomez, J. L. (1989). Late Cenozoic stress field and fracturing in the Iberian Chain and Ebro Basin (Spain). *Journal of Structural Geology*, Vol. 11, No3. pp. 285-294.
- Söderström, M. (1998). Modelling local heavy metal distribution: a study of chromium in soil and wheat at ferrochrome smelter in south-western Sweden. *Acta Agriculture Scandinavica*, Vol. 48, pp. 2-10.
- Tejero, R.; González-Casado, J.J.; Gómez-Ortiz, D. & Sánchez-Serrano, F. (2006). Insights into the “tectonic topography” of the present-day landscape of the central Iberian Peninsula (Spain). *Geomorphology*, Vol. 76, pp. 280– 294.
- Tiller, K.G. (1989). Heavy metals in soils and their environmental significance. *Advances in soil science*, Vol. 9, pp. 113-142.
- Wackernagel, H. (1998). Multivariate geostatistics: an introduction with applications. Springer, 285 pp.
- Weber, J. & Karczewska, A. (2004). Biogeochemical processes and the role of heavy metals in the soil environment. *Geoderma*, Vol. 122, pp. 105-107.

IntechOpen



## **Geochemistry - Earth's System Processes**

Edited by Dr. Dionisios Panagiotaras

ISBN 978-953-51-0586-2

Hard cover, 500 pages

**Publisher** InTech

**Published online** 02, May, 2012

**Published in print edition** May, 2012

This book brings together the knowledge from a variety of topics within the field of geochemistry. The audience for this book consists of a multitude of scientists such as physicists, geologists, technologists, petroleum engineers, volcanologists, geochemists and government agencies. The topics represented facilitate as establishing a starting point for new ideas and further contributions. An effective management of geological and environmental issues requires the understanding of recent research in minerals, soil, ores, rocks, water, sediments. The use of geostatistical and geochemical methods relies heavily on the extraction of this book. The research presented was carried out by experts and is therefore highly recommended to scientists, under- and post-graduate students who want to gain knowledge about the recent developments in geochemistry and benefit from an enhanced understanding of the dynamics of the earth's system processes.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Nikos Nanos and José Antonio Rodríguez Martín (2012). Using a Multi-Scale Geostatistical Method for the Source Identification of Heavy Metals in Soils, *Geochemistry - Earth's System Processes*, Dr. Dionisios Panagiotaras (Ed.), ISBN: 978-953-51-0586-2, InTech, Available from:  
<http://www.intechopen.com/books/geochemistry-earth-s-system-processes/using-a-multi-scale-geostatistical-method-for-the-source-identification-of-heavy-metal-in-soils>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen