

Spatial variation of trace elements in the peri-urban soil of Madrid

Antonio Vázquez de la Cueva · Ben P. Marchant ·
José Ramón Quintana · Ana de Santiago ·
Antonio López Lafuente · Richard Webster

Received: 6 November 2012 / Accepted: 13 August 2013 / Published online: 24 September 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract

Purpose The peri-urban region to the south east of Madrid contains a mixture of housing, manufacturing industry and farming, some of which disperse metals, in particular cadmium, copper, lead, and zinc, into the soil. We have mapped the concentrations of these elements and identified the major influences on their distributions.

Material and methods We sampled the topsoil at 125 sites across 1,050 km² of peri-urban land to the south east of the city on two grids, one nested inside the other. At each site, we measured the current contents of the four trace elements in the soil. We used robust geostatistical methods to model the complex spatial distributions of the data as mixtures of fixed and random effects. The empirical best linear unbiased predictor was used to map the elements. Site descriptors (lithology, land cover, cultivation, relief, erosion, and stoniness) were then included as covariates to identify significant effects on trace element concentrations.

Results and discussion The complex spatial distributions of the elements seem to arise from several sources. The concentrations generally increase from southeast to northwest, i.e., with increasing proximity to Madrid itself, the main potential source of pollution. This pattern is clear for lead and similar for copper and zinc, though with “hot spots” at or near

industrial sites. The spatial pattern of cadmium is more complex and depends on varied lithology, industry, and land use such as irrigation and cultivation. In general, the concentrations of the four elements appear to decrease with increases in stoniness and erosion, and to be largest on the valley floors.

Conclusions Robust geostatistical methods enabled us to analyze and map the complex patterns of spatial variation of trace elements in a peri-urban region of Madrid. They show that distance to the city center, lithology, manufacturing industry, and cultivation all play their parts in loading the soil with lead, copper, zinc, and cadmium. In the event, none of the metals has yet exceeded the legislative thresholds, but some concentrations are already substantially greater than would arise from natural sources, especially closest to Madrid itself.

Keywords Peri-urban · Robust geostatistics · Soil contamination · Trace metals

1 Introduction

Many big cities are expanding into their peri-urban surroundings to accommodate more and more people. The soil of these peri-urban regions is, in many instances, already polluted with potentially toxic heavy metals from city waste, industrial effluent, smoke, and even manures and fertilizers used in agriculture. Urbanization risks adding to the load and exposing a larger population to the pollution. It is important therefore to know what the current load is and to map its distribution so that (a) the risk to human health can be assessed and, if necessary, remedial action can be taken to protect people and (b) the sources of pollution can be identified and perhaps eliminated.

Numerous sources of contamination by heavy metals have been identified in peri-urban regions, and these sources act over disparate spatial scales. Point sources are factories, gas

Responsible editor: Stefan Norra

A. Vázquez de la Cueva (✉)
INIA, Forest Research Centre, Carretera de A Coruña km 7.5,
28040 Madrid, Spain
e-mail: vazquez@inia.es

B. P. Marchant · R. Webster
Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

J. R. Quintana · A. de Santiago · A. L. Lafuente
Departamento de Edafología, Facultad de Farmacia, UCM,
plaza Ramón y Cajal s/n, 28040 Madrid, Spain

works, waste deposits, and applications of compost. Cities produce large amounts of sewage, much of which contains heavy metals. The sewage is often concentrated into sludge or is composted and spread on the land nearby to minimize the cost of transport. Businelli et al. (2009), for example, found enhanced concentrations of copper (Cu), lead (Pb), and zinc (Zn) in composts added to commercial vegetable plots close to cities. Alloway (2004) reported that contamination of soil by cadmium (Cd) has been caused by the use of phosphate fertilizer containing the metal and by the disposal of ash from the combustion of coal. Copper salts are used as fungicides in vineyards, and they contribute to the large concentrations of Cu in the soil (e.g. Pietrzak and McPhail 2004; Martin et al. 2007; Saby et al. 2011). More diffuse atmospheric pollution can be attributed to traffic, industrial emissions, street dust, and the burning of fossil fuels (de Miguel et al. 1999). Nevertheless, it still tends to be local. Heavy metals from natural sources tend to vary in concentration over coarser scales according to the geological parent material. The various sources combine to form complex patterns of variation.

Cadmium, Cu, Pb, and Zn are the most widespread pollutant metals in soil. For example, Wei and Yang (2010) found that these four elements had the largest ratios of urban-to-rural concentrations in Chinese soils. El Khalil et al. 2008 found that the gradient from urban-to-suburban occupation matched the gradients in the contents of Cu and Zn in soil surrounding Marrakech. Zinc is essential for both plant and human nutrition, and Spain (BOE 1990) and some European countries (European Commission 1986) set fairly large legislative thresholds of 450 and 300 mg kg⁻¹, respectively. Although Cu is also an essential nutrient, the legislative thresholds are less (Spain 210 mg kg⁻¹; European countries 140 mg kg⁻¹) because an excess can induce deficiencies of other trace elements in livestock (Rawlins et al. 2012). Lead is biologically non-essential and toxic, but the threshold of 300 mg kg⁻¹ in Spain and other European countries probably reflects the ubiquitous presence of Pb contamination, particularly in urban areas. Cadmium is highly toxic to both humans and livestock, and the two legislative limits are set at 3 mg kg⁻¹. We might note that Switzerland has a much smaller guide value of 0.8 mg kg⁻¹ (FOEFL 1987).

We have been concerned by potentially toxic contamination of the peri-urban region to the south east of Madrid, Spain, an area of 1,050 km² into which population is spreading. A survey of the region around Madrid by de Miguel et al. (2002) found it to be the most polluted. De Miguel et al. (1999) suggested that emissions from traffic were a major source of Cu, Pb, and Zn and that construction and erosion of building materials were major sources of Cd and Zn. De Miguel et al. (1998) had also observed that the widespread use of composted sewage as fertilizer in parks and the atmospheric fallout of urban particulate material has substantially increased the concentrations of several trace elements in the urban soils of Madrid. They found that concentrations of Cu, Pb, and Zn in “undisturbed” urban

soil exceeded local natural background concentrations by factors of 2.3 to 4.0, while “compost modified” soils contained 5.3 to 8.3 times more. In the south east region, the mean concentrations were found to be 0.10 mg kg⁻¹ for Cd, 9.48 for Cu, 17.55 for Pb, and 38.53 for Zn (de Miguel et al. 2002).

We have re-sampled this south east peri-urban region to provide a more detailed picture of the pollution (a) to map the underlying distribution of the same four metals in the soil and (b) to identify the sources of the metals. We had, in addition, the hypothesis that metal concentrations increase with decreasing distance from the city center mainly because of denser traffic and industrial activities. We analyzed the data geostatistically using recently developed robust methods (Marchant et al. 2011) to quantify the spatial covariance structure and then to predict and map the concentrations of each metal across the region. These methods allow us to estimate the underlying structure largely free of the distortion that isolated hot spots of some metals from point sources might cause. We also attempt to identify attributes of the survey sites, namely lithology, land cover, erosion, stoniness, relief, and cultivation that might explain the observed variation. This has required the estimation of more than 100 linear mixed models (LMMs) by the unbiased residual maximum likelihood (REML) estimator. Such a procedure would have been impracticable until recent enhancements in computer processing power (Lark 2000).

2 Material and methods

2.1 Study region, sampling design, and analytical procedure

The region studied is a rectangle 30×35 km (1,050 km²) in the south east of the metropolitan area of Madrid. Its cover ranges from residential and industrial in the northwest to predominantly cultivated and natural vegetation in the south east. Figure 1 shows its limits within the Community of Madrid and the locations of the 125 sampled sites.

A non-preferential survey was made between during 2010 and 2011. The sampling design was based on two nested grids aligned with the Universal Transverse Mercator (UTM) projection system. The first was a 5×5 km grid. Forty eight of the 56 nodes of the grid were sampled. The second was a 1×1 km grid. Seventy four of the nodes of this grid were selected at random and sampled. On seven occasions, the selected node coincided with the 5×5 km grid, and on each occasion, the observation was made 1 km to the east. Three of the 122 observations made at grid nodes were selected at random, and for each, an additional observation was made 200 m away.

Each observation site was classified according to lithology, land cover, erosion, stoniness, relief, and cultivation. The class names and the number of elements within each class are listed in Table 1. The expected lithology and land cover were determined from a map prior to the sampling. The lithologic

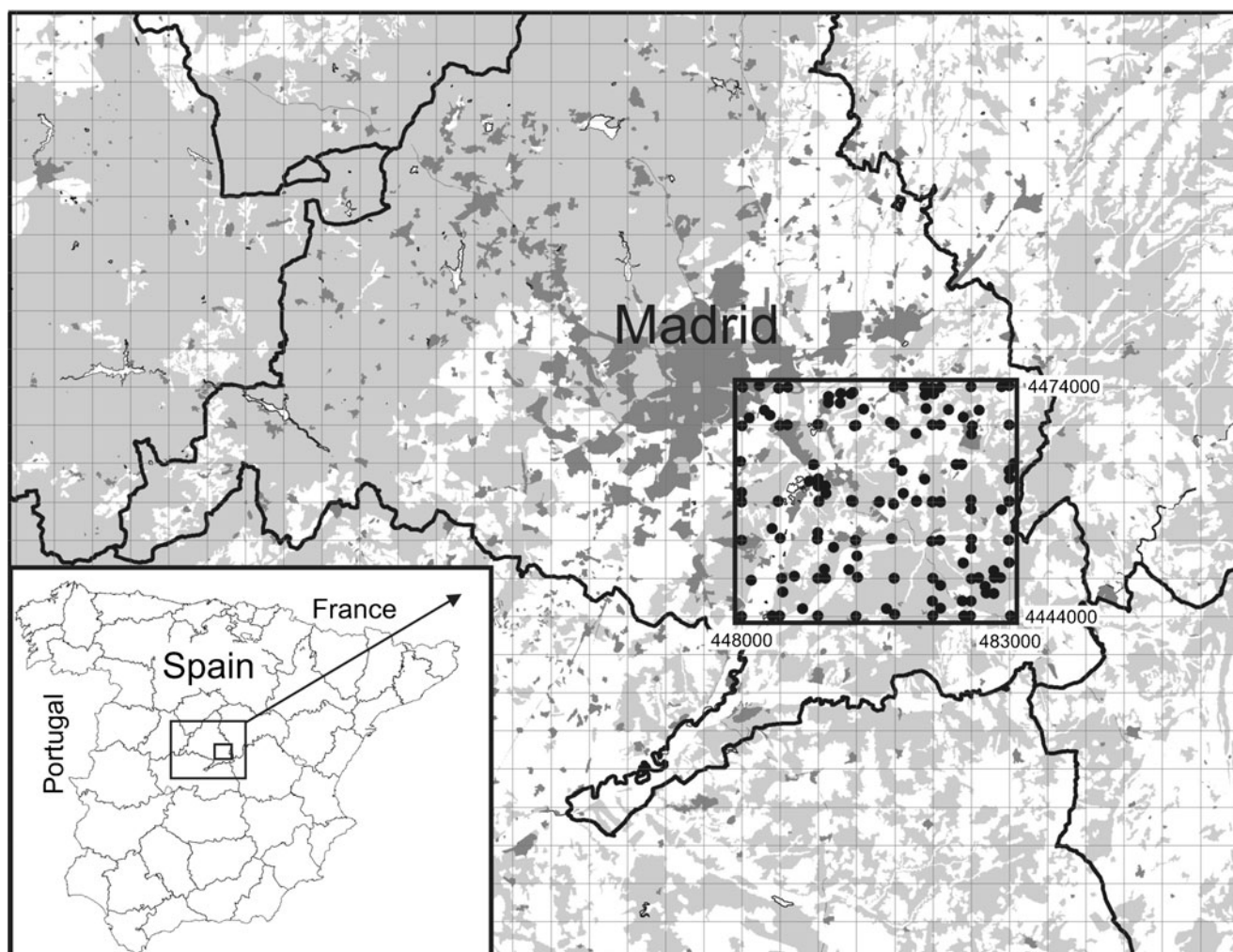


Fig. 1 Limits of the 30×35 km study region within the community of Madrid administrative boundaries and locations of the sampled points. *Dark gray* represent urban areas, *light gray* natural vegetation, and *white* cultivated areas, according to the CORINE land cover information. The overlaid 5-km grid is based in the UTM projection (zone 30)

information was based on the Instituto Geológico Minero de España (IGME) lithologic map (IGME 2012). The land cover information was derived from the CORINE Land Cover European project (NATLAN 2000) and aggregated into three classes (background of Fig. 1).

The surveyor sought to sample sites with the expected lithology and land cover attributes at the exact location of the grid node. If access to a site was denied or the site clearly did not possess the attributes expected, then an alternative location was chosen within 200 m of the specified location. If this was not possible, the location was discarded.

Soil samples were taken under homogeneous land cover. Locations were recorded with a GPS with typically less than 3 m of error. Field characteristics such as the actual lithological substrate, the land cover, the slope, and the aspect were recorded at each location. Other variables recorded were the physiographic characteristics of the location and evidence of changes in the soil horizons. Finally, each sampled location was assigned to a qualitative scale of erosion and of stoniness.

At each location, we took a composite sample of five subsamples of topsoil (0–15 cm): one in the center of the plot and the others 5 m away in the four cardinal directions. The mass of each composite sample was around 2.5 kg. This soil was air dried and later sieved to pass 2 mm. A subsample was sent to the Eurofins Analytico laboratories for the determination of the concentrations of the four metals, Cd, Cu, Pb, and Zn. The samples were subject to a full destruction with a pre-treatment based in aqua regia and microwave digestion before the analysis by inductively coupled plasma mass spectrometry and the protocol NEN-EN-ISO 17294–2. These tests are accredited by the “Dutch Accreditation Council (RvA)”.

2.2 Statistical analyses

2.2.1 Spatial prediction of trace element concentrations

The variation of the concentration of each element across the region was represented by a LMM. A LMM splits the

Table 1 Names and number of elements, *n*, of classes for each site description category

Lithology	<i>n</i>	Land cover	<i>n</i>	Cultivation	<i>n</i>
Limestone	62	Forest	6	Undisturbed	46
Conglomerate	11	Shrubland	38	Cultivated	60
Urban debris	9	Grassland	8	Urban debris	13
Gravel and sand	22	Olive trees	17	Abandoned cult.	6
Gypsum	21	Cult. Irrigated	5		
		Cult. Dryland	35		
		Urban	7		
		Vineyards	9		
Relief	<i>n</i>	Erosion	<i>n</i>	Stoniness	<i>n</i>
Flat	10	None	32	None	8
Valley bottom	23	Low	65	Low	41
Alluvial terrace	4	Medium	20	Medium	35
Slope	83	Abundant	8	Abundant	25
Top	5			Very abundant	16

variation of a property of interest into fixed effects and a spatially correlated random effect. The fixed effects are linear relationships between recorded covariates and the property. The LMM is written as

$$z = M\beta + u,$$

where *z* is a vector of *n* observations of the property of interest, *M* is a design matrix of size *n* × *q* which contains values of the *q* covariates at the observation sites, *β* is a vector of *q*-fixed effects coefficients, and *u* is realization of a second-order stationary spatially correlated random function *U* at the *n* observation sites. The random function *U* has zero mean and covariance matrix *C*. It is often assumed to have a Gaussian distribution, though this assumption is not essential. The elements of the covariance matrix were determined from authorized (Webster and Oliver 2007) parametric functions *C*(*h*) of the separation in distance and direction between pairs of observations (the lag denoted *h*) where the parameter values were estimated from the available data. It is common in the geostatistical literature for the spatial covariance of a random function to be expressed in terms of the variogram

$$\gamma(h) = \frac{1}{2}E[\{U(x)-U(x+h)\}^2],$$

where *x* = (*x*, *y*) is the location with easting *x* and northing *y*. For a second order stationary random variable

$$C(h) = C(0)-\gamma(h).$$

We had too few data to identify and estimate any anisotropy, and so we treated the lag as a separation in distance only, i.e., *h* = |*h*|. Once each LMM had been determined, it was used to predict the element *z*, at unsampled locations by

universal kriging (Webster and Oliver 2007). Universal kriging yields a prediction denoted $\hat{Z}(x_0)$ and a corresponding prediction variance $\hat{\sigma}^2(x_0)$.

The observations of elements included outliers close to local sources of pollution. These can distort the estimation of model parameters and hence, the predictions. Therefore, before estimating a LMM, we used robust geostatistical methods to identify and censor outliers. For each element, provisional LMMs were estimated by Matheron’s method of moments (Webster and Oliver 2007) and also by the robust estimators suggested by Cressie and Hawkins (1980), Dowd (1984), and Genton (1998). The distributions of all the elements were highly skewed, and the observations were transformed to their natural logarithms before these models were estimated. The fixed effects consisted of a constant and a nested nugget-plus-exponential covariance function (Webster and Oliver 2007) was assumed for the random term. Leave-one-out cross validation was done for each fitted LMM, and the squared standardized prediction error (SSPE),

$$\theta_i = \frac{[\hat{Z}(x_i)-z(x_i)]^2}{\hat{\sigma}^2(x_i)},$$

was calculated at each observation site *x_i*, *i* = 1, 2, ..., *n*. If the observed data were a realization of the fitted LMM, then the *θ_i* would be realizations of a *X*² distribution with 1 ° of freedom. The mean of this distribution is equal to 1, and the median is equal to 0.455. The mean of the *θ_i* is greatly influenced by outliers, but the median is much more robust (Lark 2002). Therefore, the LMM for which the median $\tilde{\theta}$ was closest to 0.455 was assumed to be the best representation of the underlying variation. Any observation with *θ_i* > 9 for this LMM was assumed to be an outlier, and its magnitude was reduced such

that $\theta_i=9$. An observation near to an outlier might itself be falsely classified as an outlier because of the large discrepancy between its value and that of its neighbor. Therefore, where a pair of outliers were identified within 250 m of each other, the cross-validation procedure was repeated with both potential outliers removed from the dataset (Table 2). If the recalculated θ_i at either of the two sites was less than 9, then that observation was returned to the dataset. The final LMM was fitted to these censored observations.

Various candidates for the final LMM were considered. The fixed effects were either (1) a constant, (2) linear relationships with the eastings x and northings y , or (3) a quadratic trend (i.e., linear relationships with x, y, x^2, xy and y^2). The covariance model was either pure nugget (i.e., no spatial correlation) or a nested nugget and Matérn model (Matérn 1960). This model is written:

$$c(h) = \begin{cases} c_0 + c_1 & \text{if } h = 0 \\ c_1 \left\{ \frac{1}{2^{\nu-1} \Gamma(\nu)} \left(\frac{h}{a}\right)^\nu K_\nu\left(\frac{h}{a}\right) \right\} & \text{otherwise,} \end{cases}$$

where h is the distance separating two observations, c_0 is the nugget variance, c_1 is the sill variance of the spatially correlated component, a is a distance parameter, ν is a smoothness parameter which yields flexibility in how the covariance function behaves for small h , $\Gamma(\nu)$ is the Gamma function and K_ν is a modified Bessel function of the second kind of order ν (Abramowitz and Stegun 1972). If $\nu=0.5$ then the Matérn model is equivalent to the exponential model (Webster and Oliver 2007). We transformed the observed values by the Box–Cox method (Diggle and Ribeiro 2007) so that they were consistent with the assumed Gaussian random effects of the LMM. The Box–Cox transform is

$$z^* = \begin{cases} \frac{(z^\lambda - 1)}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(z) & \text{if } \lambda = 0 \end{cases}$$

where λ is a parameter.

The parameters of the LMM, $\alpha = (c_0, c_1, a, \nu, \beta, \lambda)$, were estimated by the maximum likelihood (ML) estimator described by Diggle and Ribeiro (2007). This selects the model parameters that maximize the likelihood that the data arose from the model. The different fitted candidate models were compared by calculation of the Akaike Information Criterion (AIC; Akaike 1973):

$$AIC = -2L(\hat{\alpha}|z) + 2p,$$

where p is the number of parameters in the LMM and $L(\hat{\alpha}|z)$ is the log likelihood that the observed data arose from the LMM with estimated parameters $\hat{\alpha}$. The model with the

smallest AIC strikes the best balance between complexity (number of parameters) and closeness of fit to the data (likelihood). There is a small bias in the ML parameter estimates of a LMM when the fixed effects are not constant. Therefore, once the model with the smallest AIC had been found, it was re-fitted by the REML estimator (Diggle and Ribeiro 2007) which minimizes this bias. Then, this model was used to predict the concentration across the region by universal trans-Gaussian kriging (Diggle and Ribeiro 2007).

2.2.2 Identification of major influences on trace element concentrations

A stepwise regression algorithm incorporating likelihood ratio tests was used to determine site descriptor classes that had significant effects on concentrations. We note that these descriptors could not be used to predict the concentrations across the region because they were known only at the sampling sites. The 33 classes of the six location classifications (see Table 1) were treated as indicator variables which described the presence or absence of the class at each site. The eastings and northings, x and y formed two further potential covariates. Some of these potential covariates are strongly correlated. This means that the order in which covariates are added to the LMM can influence whether or not a particular covariate is included because a correlated fixed effect might already be included. Therefore, an automated iterative algorithm was implemented.

Initially, a LMM with constant mean and nested nugget and Matérn covariance structure was fitted to the element being investigated. Further models were then fitted where the fixed effects contained an additional covariate. Each of the potential covariates was considered in turn. The covariate that increased the likelihood by the largest amount was added to the fixed effects if the likelihood ratio test showed that the increase was significant at the $p=0.15$ level. This process was the forward selection step. The addition of a covariate to a model might have meant that one of the covariates already included in the fixed effects was no longer required. Therefore a backward elimination step was then applied. In this step, the model was re-fitted with one covariate removed from the fixed effects. This was repeated for all of the covariates in the fixed effects, and the covariate removal that led to the smallest reduction in the likelihood was found. If this reduction in the likelihood was not significant at the $p=0.15$ level, then the covariate was removed from the model and a further backward elimination step was made. Otherwise, a further forward selection step was made. The algorithm continued until neither a forward selection nor a backward elimination step led to a change in the fixed effects.

Such a stepwise procedure does not necessarily determine the best covariates to include in the fixed effects of a model. Other fixed effects might have performed similarly, and when the aim of a study is to determine the best model for predictive

Table 2 Summary and cross-validation statistics

	Cd	Cu	Pb	Zn
Mean/mg kg ⁻¹	0.14	12.95	27.95	42.75
Median/mg kg ⁻¹	0.11	9.50	17.00	40.00
Std dev./mg kg ⁻¹	0.08	21.33	98.73	22.04
Min/mg kg ⁻¹	0.01	3.03	5.60	13.00
Max/mg kg ⁻¹	0.48	240.00	1100.00	180.00
Skew	1.71	9.80	10.43	2.91
Skew of $\ln(z)$	-1.11	1.93	3.72	0.38
Matheron $\tilde{\theta}$	0.25	0.25	0.20	0.39
Best estimator	C-H*	Dowd	Genton	Matheron
Robust $\tilde{\theta}$	0.42	0.49	0.41	0.39
Large outliers	1	1	3	1
Small outliers	2	1	1	0
λ	-0.02	-0.26	-0.08	-0.07

*Cressie-Hawkins

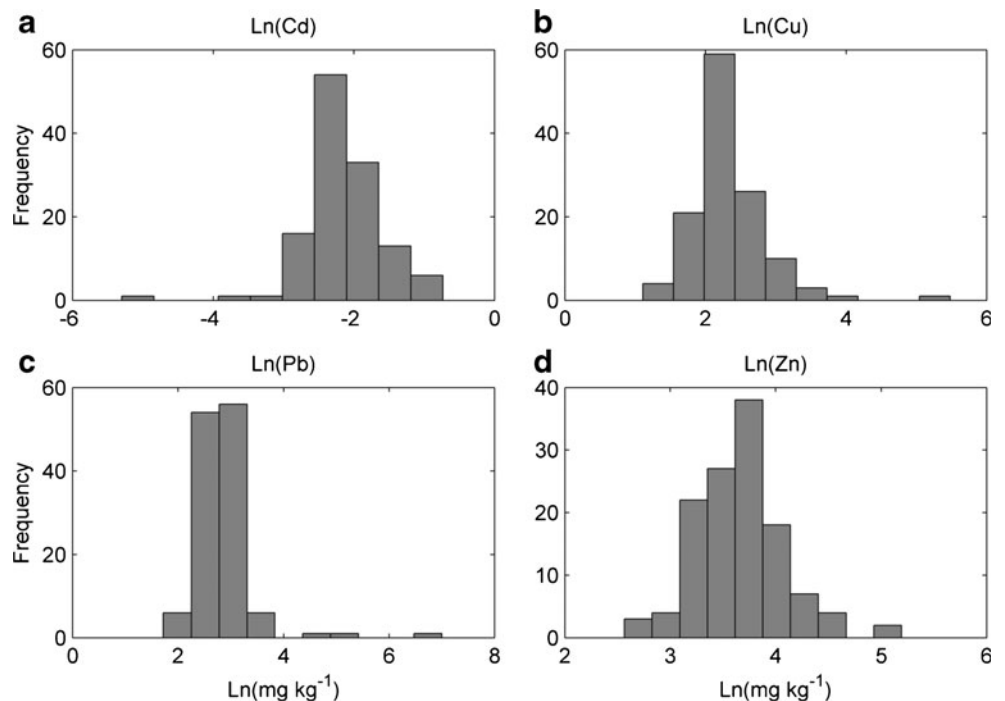
purposes, it is important that one takes into account one’s knowledge of the covariates. However, the procedure did indicate some covariates that influence concentrations in an objective manner.

3 Results

Table 2 summarizes the statistics for the observed concentrations. The frequency distributions of all elements are positively skewed: they have skewness coefficients greater than 1. This skew is reduced by a log transform (Fig. 2). In the case of cadmium, the log transform produced a strong negative skew.

However, the histogram of log Cd suggests that this negative skew might be the result of three unusually small measurements. When the variograms fitted to the concentrations of Cd, Cu, and Pb by Matheron’s estimator were cross-validated, the median SSPEs ($\tilde{\theta}$) were less than 0.3. This indicates that outliers are present amongst the observations of these properties. When robust variogram estimators were used, the median SSPEs were between 0.4 and 0.5. Different robust variogram estimators performed best on each of these three properties. The median SSPE for the Matheron variogram of Zn was 0.39, and this model fitted the data better than those estimated by robust methods.

Fig. 2 Histograms of log transformed observations of trace elements within study region



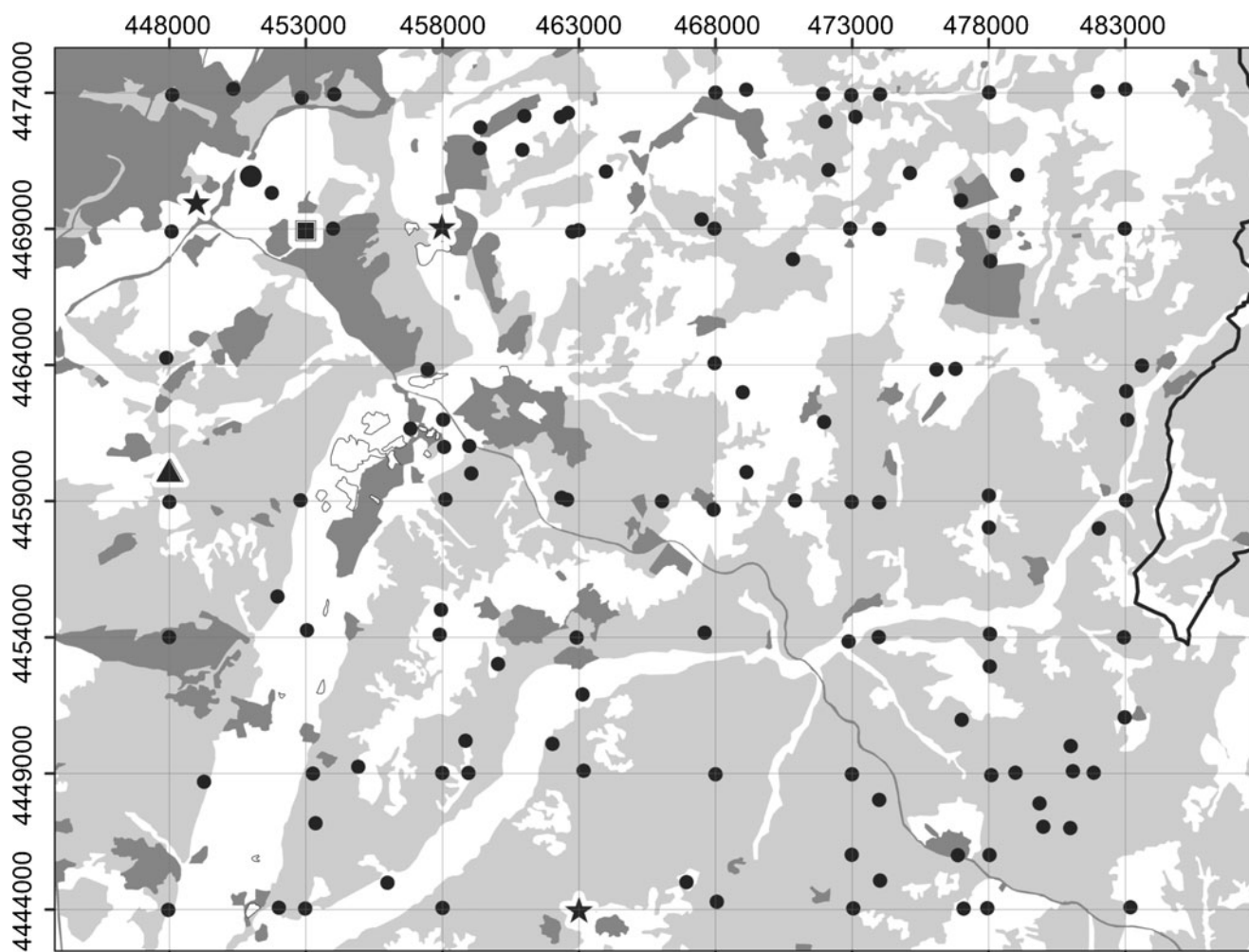


Fig. 3 Sample design and location of large outlying observations of Cd (black square), Cu (big disc), Pb (black star), and Zn (black triangle). Dark gray represent urban areas, light gray natural vegetation, and white cultivated areas, according to the CORINE land cover information. The overlaid 5-km grid is based in the UTM projection (zone 30)

Six large outliers (one of Cd, one of Cu, three of Pb, and one of Zn) were identified and censored. They are all in the urban northwest corner of the region (Fig. 3). A further four unusually small outliers were identified. We considered these to be less relevant than the large outliers because they required only small censoring adjustment in the untransformed units (mg kg^{-1}).

The calculated AIC values (Table 3) showed that the best fitting LMMs for each of the four censored elements included a linear trend in the eastings and northings. The models for Cd, Cu, and Zn included a Matérn term in the random effect, whereas there was no evidence of spatial correlation in the model for Pb. The REML-estimated variograms for the random effects of these models are shown in Fig. 4. The variograms for Cu and Zn are very similar and have a smaller nugget-to-sill ratio and smaller range than the model for Cd.

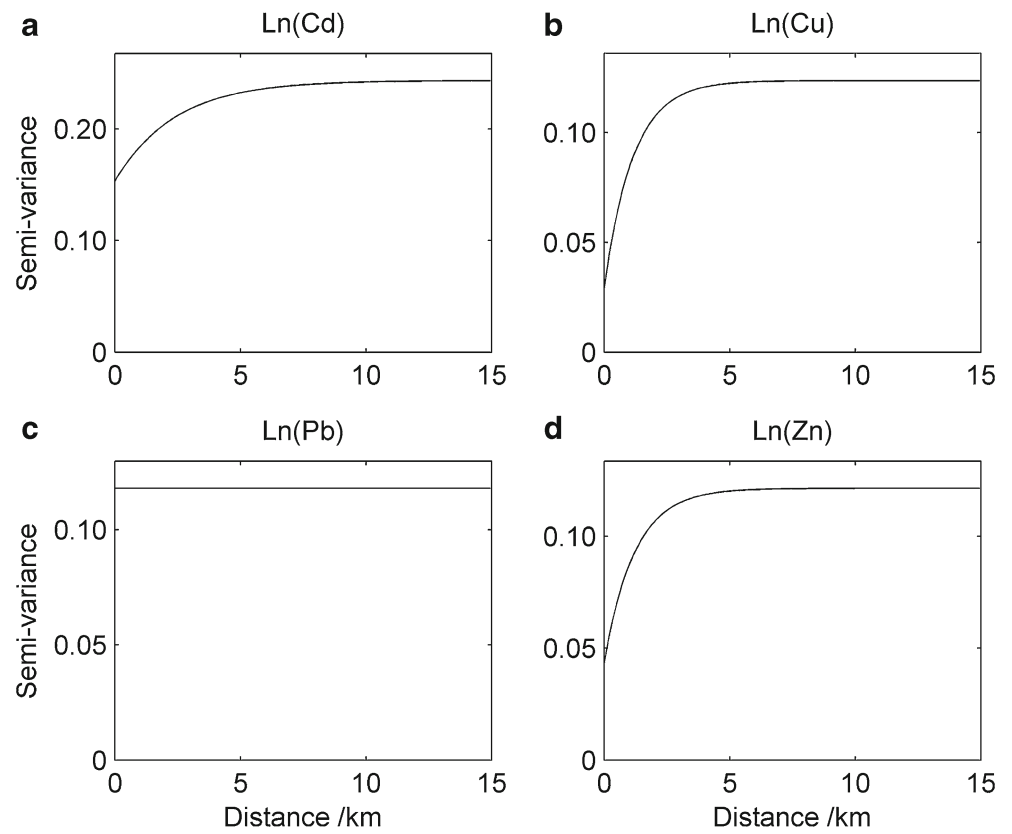
The predictions (see the maps in Fig. 5) of Cu, Pb, and Zn, decrease in concentrations from the urban northwest corner to

the rural south east corner. For Cd, the trend is more from the west to east. The patterns of variation for Cu and Zn are fairly similar with several fairly small areas of contamination. The variation of Cd is smoother. The absence of spatial correlation in the LMM of Pb means that the spatial trend is the only cause of variation.

Table 3 AIC for models fitted by maximum likelihood. The smallest AIC value for each metal is shown in bold

Trend	Variogram	Cd	Cu	Pb	Zn
Constant	Matérn	-160.57	-180.91	-148.63	-178.76
Linear	Matérn	-166.54	-194.03	-154.57	-194.20
Quadratic	Matérn	-164.55	-191.58	-149.71	-190.99
Constant	Nugget	-157.67	-169.79	-149.64	-159.16
Linear	Nugget	-164.55	-191.31	-155.86	-192.63
Quadratic	Nugget	-164.76	-190.18	-150.87	-189.50

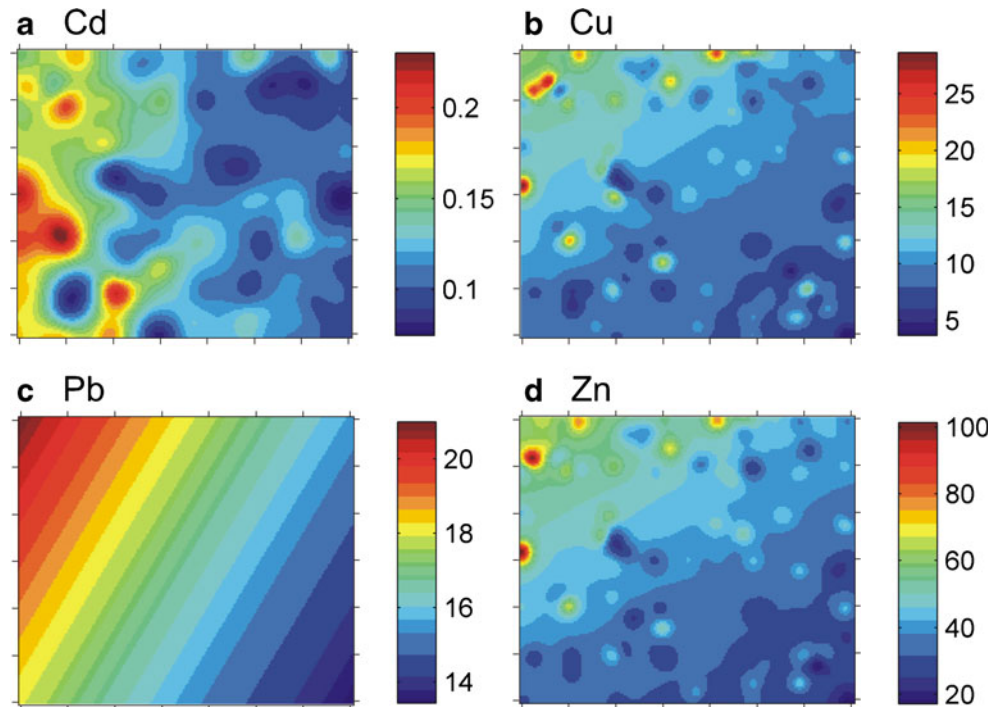
Fig. 4 Estimated variograms for best fitting models



The most effective indicators to include in the fixed effects are listed in Table 4. Membership of some of the classes has a similar effect on the concentration of all four trace elements.

For example, concentrations appear to decrease with increases in stoniness and erosion, and valley floors have large concentrations. Copper, Pb, and Zn all have small concentrations on

Fig. 5 Spatial prediction of metal concentrations (in milligrams per kilogram) across the study region



gravel and sand. Land cover classes feature fewer times. This might be an artifact caused by the large number of these classes. There are some differences between the four elements. More lithology classes are included in the model for Cd than for the others, and gypsum is the first class to be included for it. Stoniness appears to be important for Zn.

4 Discussion and conclusions

There are some similarities in the spatial patterns of the four trace elements we have studied. The concentrations of all are generally largest in the northwest of the region. This result is likely to be caused by the fallout of airborne particles originating in urban Madrid further to the northwest and from road traffic on the highways across this part of the region (de Miguel et al. 1999; Manta et al. 2002; Li et al. 2004). The concentrations of all four elements appear to decrease with increases in stoniness and erosion, and to be largest in the fine sediments on the valley floors. This accords with general knowledge: trace metals tend to be adsorbed predominantly on clays (Sillanpää 1972), and the fine fraction of soil mobilized by erosion accumulates with adsorbed metals on valley floors.

The spatial patterns of Cu and Zn, in particular, are similar. Other authors (Adachi and Tainosho 2004; Li et al. 2004; Yang et al. 2011) have described how these metals in car components and lubricants can be emitted into the atmosphere as small particles and then be deposited on the ground. Such emissions are probably the major source of contamination in the northwest of the region. Hot spots of Cu and Zn are evident further from Madrid. These coincide with industrial sites such as a large urban waste incineration plant, a sepiolite mining installation, and other construction treatment plants. Morton-Bermea et al. (2009) observed a similar pattern of Cu and Zn contamination in Mexico City. There, the underlying pattern was attributed to emissions from motor traffic with local hot spots close to industrial sites.

The first factor in the best-fitting LMM of Cu (see Table 4) was undisturbed (non-cultivated) soils, which suggest that the concentrations on cultivated soils are larger than elsewhere. This might be because Cu salts have been used as fungicides, as other investigators have found (Brun et al. 1998; Micó et al. 2007).

The map of lead shows only a decreasing trend from the northwest. Lead used to be added to petrol (the practice ceased in 1999; Xia et al. 2011), and emissions in exhaust are likely to have been the main source of lead in the region. No other features can be seen in the map because the random effects were not spatially correlated.

The variation of Cd across the region seems to depend on several factors. Table 4 indicates the importance of lithology. Previous studies (e.g. Atteia et al. 1995; Marchant et al. 2010)

have also found that Cd in the parent material of the soil contributes substantially to that in the soil itself. In the region of our study, Cd is most concentrated in the two main areas of gypsum. The Cd concentrations are also largely close to the urban center which includes river valleys where intensive and irrigated agriculture is practiced. Other authors have found a close relation between Cd concentrations and intensive agriculture caused by the use of phosphate fertilizers and irrigation water contaminated by urban or industrial effluents (Hu et al. 2006; Chen et al. 2009; Acosta et al. 2011). Cadmium hotspots are also evidently close to the plant for incinerating urban waste.

The four elements vary over disparate spatial scales, and we presume they are affected by several factors. Local sources of contamination such as industrial emissions, construction materials, and uncontrolled waste deposits produced outliers. The natural variation seems to be related largely to lithology (Adriano 1986; Kabata-Pendias and Pendias 2001). There is

Table 4 Site descriptor classes included in the best-fitting LMM for each metal and effect of membership of class upon metal concentration

Cd			Cu		
Category	Class	Effect	Category	Class	Effect
Lithology	Gypsum	+	Cultivation	Undisturbed	–
Relief	Bottom	+	Stoniness	None	+
Erosion	Medium	–	Spatial	Northing	+
Spatial	Easting	–	Relief	Top	–
Lithology	Conglomerate	–	Lithology	Gravel and sand	–
Relief	Alluvial terrace	–	Land cover	Irrigated cult	+
Lithology	Limestone	–	Spatial	Easting	–
Spatial	Northing	+	Relief	Slope	–
Land cover	Shrub land	+	Cultivation	Cultivated	–
Cultivation	Cultivated	–	Land cover	Grassland	–
Erosion	None	+	Relief	Alluvial terrace	–
Pb			Zn		
Category	Class	Effect	Category	Class	Effect
Spatial	Easting	–	Stoniness	None	+
Erosion	None	+	Stoniness	Low	+
Land cover	Unirrigated cult	–	Spatial	Northing	+
Stoniness	Abundant	–	Relief	Flat	+
Relief	Alluvial terrace	–	Spatial	Easting	–
Cultivation	Abandoned	+	Lithology	Gravel and sand	–
Land cover	Irrigated cult	+	Relief	Valley bottom	+
Lithology	Gravel/sand	–	Relief	Slope	+
Stoniness	Medium	–	Erosion	Medium	–
Spatial	Northing	+	Cultivation	Abandoned	+
			Erosion	Abundant	–

great lithological diversity in the sedimentary rocks and evaporites. It is for this reason that standard linear geostatistical models were not suitable for analyzing the data and that robust methods were required to identify outliers and reduce their influence. Similarly, the determination of the factors controlling the variation was far from trivial. We iteratively built up a LMM for each element using an automated stepwise algorithm. The results of such an approach should be interpreted with caution. The site descriptors considered were highly correlated, and therefore the inclusion of one factor within the model might mask the effect of another. In contrast to previous studies, there was little evidence that composts were a substantial source of the metals. Indeed, when cultivation does appear in the LMMs in Table 4, it has a negative effect. This might be only relative to soil containing urban debris. Despite these concerns, some general causes of variation can be discerned from the fitted LMMs such as the effects of erosion, relief, and stoniness discussed above.

The concentrations of the four elements were generally less than the legal limits established for agriculture by the European Commission (1986) and Spain (BOE 1990). Only one observation of Pb and one of Cu exceeded these limits. The mean values of all four metals were less than those reported in other peri-urban areas in towns with longer industrial histories (Kelly et al. 1996; Frangi and Richard 1997; Manta et al. 2002; Imperato et al. 2003). The contamination in Madrid has occurred relatively recently and therefore, the enrichment in metals of the soil is more similar to that found in towns of developing countries with more recent industrialization (Hu et al. 2006; Yang et al. 2009; Xia et al. 2011). Note that de Miguel et al. (1998) found larger concentrations in the centre of Madrid. Nevertheless, the concentrations in this peri-urban region exceed the natural concentrations, and we should be aware of further contamination which might lead to toxic quantities entering the food chain and eventually causing ill health.

Acknowledgments We thank the CARESOIL_CM network (Project founded by the Comunidad Autónoma de Madrid, P2009/AMB-1648). B.P. Marchant's contribution was funded by the Biotechnology and Biological Sciences Research Council (BBSRC).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Abramowitz M, Stegun IA (eds) (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover Publications, New York
- Acosta JA, Faz A, Martínez-Martínez S, Arocena JM (2011) Enrichment of metals in soils subjected to different land uses in a typical Mediterranean environment (Murcia City, southeast Spain). *Appl Geochem* 26:405–414
- Adachi K, Tainosho Y (2004) Characterization of heavy metal particles embedded in tire dust. *Environ Int* 30:1009–1017
- Adriano DC (1986) Trace elements in the terrestrial environment. Springer, New York
- Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csáki F (eds) Second International Symposium on Information Theory. Akadémiai Kiadó, Budapest, pp 267–281
- Alloway BJ (2004) Contamination of soils in domestic gardens and allotments: a brief overview. *Land Contam Reclam* 12:179–187
- Attea O, Thélin P, Pfeifer HR, Dubois JP, Hunziker JC (1995) A search for the origin of cadmium in the soil of the Swiss Jura. *Geoderma* 68:149–172
- BOE (1990) Real Decreto 1310/1990 de 29 de octubre por el que se regula la utilización de los lodos de depuración en el sector agrario. Boletín Oficial del Estado, núm. 262 de 1 de noviembre de 1990, pp 32339–32340, BOE-A-1990-26490
- Brun LA, Maillot J, Richarte J, Herrmann P, Remy JC (1998) Relationships between extractable copper, soil properties and copper uptake by wild plants in vineyard soils. *Environ Pollut* 102:151–161
- Businelli D, Massaccesi L, Said-Pullicino D, Gigliotti G (2009) Long-term distribution, mobility and plant availability of compost-derived heavy metals in a landfill covering soil. *Sci Total Environ* 407:1426–1435
- Chen T, Liu X, Li X, Zhao K, Zhang J, Xu J, Shi J, Dahlgren RA (2009) Heavy metal sources identification and sampling uncertainty analysis in a field-scale vegetable soil of Hangzhou, China. *Environ Pollut* 157:1003–1010
- Cressie N, Hawkins D (1980) Robust estimation of the variogram. *J Int Ass Math Geol* 12:115–125
- De Miguel E, de Grado UMJ, Llamas JF, Martín-Dorado A, Mazadiego LF (1998) The overlooked contribution of compost application to the trace element load in the urban soil of Madrid (Spain). *Sci Total Environ* 215:113–122
- De Miguel E, Llamas JF, Chacón E, Mazadiego LF (1999) Sources and pathways of trace elements in urban environments: a multi-elemental qualitative approach. *Sci Total Environ* 235:355–357
- De Miguel E et al (2002) In: Baretino D, Callaba A (eds) Determinación de niveles de fondo y niveles de referencia de metales pesados y otros elementos traza en suelos de la Comunidad de Madrid. Instituto Geológico y Minero de España, Madrid
- Diggle PJ, Ribeiro PJ (2007) Model-based geostatistics. Springer, New York
- Dowd PA (1984) The variogram and kriging: robust and resistant estimators. In: Verly G, David M, Journé AG, Marechal A (eds) Geostatistics for Natural Resources Characterization. Part 1. Reidel, Dordrecht, pp 91–106
- El Khalil H, Schwartz C, Elhamiani O, Kubiniok J, Morel JL, Boularbah A (2008) Contribution of technic materials to the mobile fraction of metals in urban soils in Marrakech (Morocco). *J Soils Sediments* 8:17–22
- European Commission (1986) European Commission Council Directive 86/278/EEC of 12 June 1986 on the protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture. *Off J Eur Comm L* 181:6–12
- FOEFL (1987) Commentary on the ordinance relating to pollutants in soil (VSBo of June 1986). FOEFL (Federal Office of Environment, Forest and Landscape), Bern
- Frangi JP, Richard D (1997) Heavy metal soil pollution cartography in northern France. *Sci Total Environ* 205:71–79
- Genton MG (1998) Highly robust variogram estimation. *Math Geol* 30:213–221
- Hu KL, Zhang FR, Li H, Huang F, Li BG (2006) spatial patterns of soil heavy metals in urban–rural transition zone of Beijing. *Pedosphere* 16:690–698

- IGME (2012) Instituto Geológico y Minero de España. Cartografía Geológica MAGNA (in digital format). <http://www.igme.es>. Accessed May 2012
- Imperato M, Adamo P, Naimo D, Arienzo M, Stanzione D, Violante P (2003) Spatial distribution of heavy metals in urban soils of Naples city (Italy). *Environ Pollut* 124:247–256
- Kabata-Pendias A, Pendias H (2001) Trace elements in soils and plants, 2nd edn. CRC Press, Boca Raton, Florida
- Kelly J, Thornton I, Simpson PR (1996) Urban geochemistry: a study of the influence of anthropogenic activity on the heavy metal content of soils in traditionally industrial and non-industrial areas of Britain. *Appl Geochem* 11:363–370
- Lark RM (2000) Estimating variograms of soil properties by the method-of-moments and maximum likelihood. *Eur J Soil Sci* 51:717–728
- Lark RM (2002) Modelling complex soil properties as contaminated regionalized variables. *Geoderma* 106:1731–1790
- Li XD, Lee SL, Wong SC, Shi WZ, Thornton L (2004) The study of metal contamination in urban soil of Hong Kong using a GIS-based approach. *Environ Pollut* 129:113–124
- Manta DS, Angelone M, Bellanca A, Neri R, Sprovieri M (2002) Heavy metals in urban soils: a case study from the city of Palermo (Sicily), Italy. *Sci Total Environ* 300:229–243
- Marchant BP, Saby NPA, Lark RM, Bellamy PH, Jolivet CC, Arrouays D (2010) Robust analysis of soil properties at the national scale: cadmium content of French soils. *Eur J Soil Sci* 61:144–152
- Marchant BP, Tye AM, Rawlins BG (2011) The assessment of point-source and diffuse soil metal pollution using robust geostatistical methods: a case study in Swansea (Wales, UK). *Eur J Soil Sci* 62: 359–370
- Martín JAR, de la Cueva AV, Corbi JMG, López-Arias M (2007) Factors controlling the spatial variability of copper in topsoils of the north-eastern region of the Iberian Peninsula, Spain. *Water Air Soil Pollut* 186:311–321
- Matérn B (1960) Spatial variation. *Meddelanden från Statens Skogsforskningsinstitut*, 49, No 5, 2nd edn. Springer, New York, Lecture Notes in Statistics No. 36
- Micó C, Peris M, Recatala L, Sánchez J (2007) Baselines values for heavy metals in agricultural soils in an European Mediterranean region. *Sci Total Environ* 378:13–17
- Morton-Bermea O, Hernández-Álvarez E, González-Hernández G, Romero F, Lozano R, Beramendi-Orosco LE (2009) Assessment of heavy metal pollution in urban topsoils from the metropolitan area of Mexico City. *J Geochem Explor* 101:218–224
- NATLAN (2000) CLC 1990, CORINE Land Cover 250m. European Environment Agency, Copenhagen. <http://www.eea.europa.eu/data-and-maps>. Accessed May 2011
- Pietrzak U, McPhail DC (2004) Copper accumulation, distribution and fractionation in vineyard soils of Victoria, Australia. *Geoderma* 122: 151–166
- Rawlins BG, McGrath SP, Scheib AJ, Breward, N Cave, M, Lister TR, Ingham, M, Gowing, C, Carter, S (2012) The advanced soil geochemistry atlas of England and Wales. British Geological Survey, Keyworth. www.bgs.ac.uk/gbase/advsoilatlasEW.html. Accessed Jan 2012
- Saby NPA, Marchant BP, Lark RM, Jolivet CC, Arrouays D (2011) Robust geostatistical prediction of trace elements across France. *Geoderma* 162:303–311
- Sillanpää M (1972) Trace elements in soils and agriculture. *Soils Bulletin* 17. FAO, Rome, 67 pp
- Webster R, Oliver MA (2007) *Geostatistics for environmental scientists*, 2nd edn. John Wiley and Sons, Chichester
- Wei BG, Yang LS (2010) A review of heavy metal contaminations in urban soils, urban road dusts and agricultural soils from China. *Microchem J* 94:99–107
- Xia X, Chen X, Liu R, Liu H (2011) Heavy metals in urban soils with various types of land use in Beijing, China. *J Hazard Mater* 186: 2043–2050
- Yang P, Mao R, Shao H, Gao Y (2009) An investigation on the distribution of eight hazardous heavy metals in the suburban farmland of China. *J Hazard Mater* 167:1246–1251
- Yang Z, Lu W, Long Y, Bao X, Yang Q (2011) Assessment of heavy metals contamination in urban topsoil from Changchun City, China. *J Geochem Explor* 108:27–38