

# Evaluation of the atmospheric deposition of heavy metals with biosensors in the region of Lisbon

Ricardo Gomes Junqueira Morais

## Abstract

The main purpose of this work was to evaluate the atmospheric deposition of 10 heavy metals (As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Zn) in the region of Lisbon and the south margin of the Tagus river, based on the deposition values measured on 28 biomonitoring samples. The biomonitor used was the lichen species *Xanthoria parietina* (L.) Th.Fr. Firstly, the average concentrations of heavy metals on the lichen tissues were compared to those of other studies, leading to the conclusion that, in general, the average levels of atmospheric deposition in Lisbon and the south margin of the Tagus are low and/or normal, with the possible exception of As. The village of Paio Pires (in Seixal), the industrial area of Barreiro, and, to a lesser extent, Avenida da Liberdade, are the places with the highest deposition levels of Cd, Co, Cu, Fe, Mn, Ni, Pb e Zn, while the urban park of Vale Fundão (Chelas), Sobreda (Almada) and Avenida da Liberdade have the highest levels of As and Cr.

Secondly, the maps of concentrations of those 10 heavy metals for the study area were generated, using the Direct Sequential Co-simulation algorithm, with two different approaches. The first one consists in simulating the maps of one element conditioned to the simulation of another element, while the second approach, which is used for the first time in this thesis, consists of simulating the maps of the elements conditioned to one of the Principal Components (a result of the Principal Components Analysis technique). For each of these approaches, 50 maps were simulated, for each heavy metal, and the histograms and variograms were calculated for a fraction of those maps, to evaluate the reproducibility of these statistics given by the two approaches. It was concluded that both methods produced equally satisfying results, with some limitations given by the nature and quantity of the original data. The maps of average values for each 50 simulations were also calculated, along with the maps of standard deviations, for the study area, and it was concluded that the simulation is more confident (has less variability, then less uncertainty) in the areas that are near the samples, and gives great importance to the value of those samples when simulating the nodes around them. Finally it was concluded that the simulation based on the Principal Components produces concentration maps that have more slightly symmetrical distributions and with more homogeneous variability than the maps produced by the other approach, when the Principal Components synthesize a great number of variables.

**Key-words:** biomonitoring; atmospheric deposition; heavy metals; Direct Sequential Co-simulation; Principal Components; Lisbon and the south margin of the Tagus river

## 1. Introduction

Biomonitoring are organisms that allow the monitoring of changes in the environment, through the quantification of different parameters in their tissues. Lichens, in particular, are often used as biomonitors of air quality, because of their ubiquitous presence in the environment, because they are relatively easy to collect and analyze, they do not use their substratum as a source of nutrients, and they do not possess outward protection structures nor structures specialized in obtaining water and nutrients, and so direct changes occur between the lichen surface and the atmosphere (Bargagli & Mikhailova, 2002).

Many studies involving biomonitoring evaluate the heavy metal concentrations in the atmospheric deposition, because they are elements with significant impacts on ecosystems and on human health, and because air is an important path of dispersion and exposure to these pollutants. Metals like Pb, Cd or As pose a great threat to human health, being the cause for many different diseases, as reported by several studies. Pb is responsible for having neurotoxic effects on children, while exposure to Cd may bring about kidney problems and bone related diseases, and As may be the cause for different types of cancer (Järup, 2003). Heavy metals may have a natural origin, for example Fe, Ni and Cr, whose presence in the atmosphere may be a result of the lithology, through the erosion and suspension of soil particles, but the most important sources are of human origin. Some activities that contribute to high emission levels of these metals are mining, the production of metal alloys, refining and burning of fossil fuels, the wear of automobile components, such as tires and brakes, etc.

Geostatistics is a field of statistics that is dedicated to the characterization of the spatial or spatio-temporal structure of phenomena from which there is limited knowledge (based on a certain number of samples). The techniques available in geostatistics allow the estimation/simulation of values of a certain attribute for areas where there is no information available, based on a usually small number of samples of that attribute. This may apply to the characterization of ore deposits, oil reservoirs, indexes of tree vitality, atmospheric deposition as assessed by biomonitors, etc., and allows the quantification of the uncertainty associated with the estimation/simulation process. There are many studies that use geostatistics in the field of biomonitoring heavy metal pollution, such as Figueira et al. (2002) and Reis et al. (2010).

Alternatively to estimation procedures like kriging, it is common to use stochastic simulation to obtain the values for non-sampled areas, as it accounts for the variability (uncertainty) associated with the attribute under consideration. Simulation algorithms give different scenarios for the maps of values of that attribute, which are possible realizations of the spatial distribution in that area and have all the same probability of occurrence. Direct Sequential Simulation (Soares, 2001), simplified as DSS, is a simulation method that is able to reproduce the statistics of the samples, namely the histogram and the variogram, and, unlike some other simulation algorithms, does not need to transform the original data. When a variable is scarcely sampled, and is highly correlated to another variable, for example some kind of environmental factor, for which there is a large quantity of data available, in the study area, it may be useful to simulate the first variable conditioned to the second one, in a process called co-simulation. This can be applied with the DSS, in a process called Direct Sequential Co-simulation. This has been used, for example, by Franco et al. (2006).

The main purpose of this work was to evaluate the atmospheric deposition of 10 heavy metals (As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Zn) in the region of Lisbon and the south margin of the Tagus, using data from a biomonitor species, the lichen *Xanthoria parietina*. The average values of concentrations in the lichen tissues will be compared to those found in other studies, and maps of concentrations for the study area will be simulated, using the co-DSS algorithm. The simulation step will be done with two approaches: co-simulation conditioned to the concentration maps of another metal, and co-simulation conditioned to the Principal Component with the highest correlation to each metal. The latter approach is a new procedure in terms of the simulation practice, and requires the use of the PCA (Principal Components Analysis) technique, which is a method for data analysis, that transforms a set of  $n$  original variables into a set of  $k$  new variables, which synthesize the information and relationships found in the original ones, with  $k < n$ . Finally, a comparison between the two simulation approaches is intended, using the variograms and histograms of the simulations, and also the maps with the simulated values, and with the average and standard deviation values.

## 2. Methods

Firstly, the concentrations of the 10 heavy metals (As, Cd, Co, Cr, Cu, Fe, Mn, Ni, Pb, Zn) in the tissues of the 28 biomonitor samples (Figure 1) were analyzed individually, using some basic descriptive statistics like the average, standard deviation, etc., and the main hotspots of atmospheric

deposition concentrations were identified. The average values found in the study area were compared to a number of studies using the same lichen species, in order to evaluate if the concentration levels in the atmospheric deposition were acceptable. Secondly, the Principal Components Analysis (PCA) technique was used to synthesize the main relationships between the concentrations in different metals. A set of  $k$  Principal Components (PC) whose associated eigen values were higher than 1, and that could explain most of the variability (at least 80%) of the original data was retained.

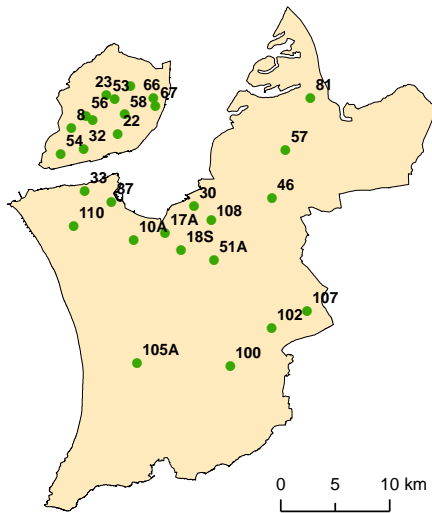


Figure 1 – Geographical location of the sampling places and sample identification.

Then, it was carried out an analysis of the spatial structure of each variable, with the help of the variogram,  $\gamma(\mathbf{h})$ , a statistical tool that describes the average dissimilarity between samples separated by a vector  $\mathbf{h}$ . The variograms for the directions  $45^\circ$  (NE/SW) and  $-45^\circ$  (NW/SE) were calculated, and a theoretical model was fitted to them. The functions used to model variograms are usually simple ones, with few parameters, like the spherical model, which was used in this case, and is defined as follows:

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[ 1.5 \frac{h}{a} - 0.5 \left( \frac{h}{a} \right)^3 \right] & \text{for } h \leq a \\ C & \text{for } h > a \end{cases}$$

The parameters of this model are the sill,  $C$ , the nugget effect,  $C_0$ , which corresponds to small scale variability, and is given by the y-intercept, and the range,  $a$ , which is the  $h$  value for which the variogram reaches the sill. The sum  $C+C_0$  is normally equal to the variance.

In order to enrich the simulation process, the correlations between the metal concentrations and some auxiliary variables that could be used for co-simulation were studied. Variables such as land use, type of geologic formations, energy consumption, population and distance to roads were tested. For the first two, the proportions of different land uses and types of geologic formations were calculated in a circular area with a 500 meter radius around each sample, using ArcMap10™ (Esri®), and the correlations between the proportions and the metal concentrations were obtained. For the other variables, the value for each sample was determined, using information from population census, the national statistics institute (INE), etc., and the correlations were calculated.

As the metal concentrations were not significantly correlated to the auxiliary variables, the simulation step was done independently of these variables. Two approaches were used: simulation of 50 maps of concentrations of one metal, based on the previous simulation of another metal, which has the highest correlation with the first one; simulation based on the PC with the highest correlation with the desired metal. For the first method, the variables were set into two groups, based on the correlation coefficients. As and Cr were set in one group, while the other 8 (Cd, Co, Cu, Fe, Mn, Ni, Pb, Zn) made up the other. For each group, one of the metals was chosen to start the simulation process, done by DSS, based on the characteristics of the variogram structure. As and Fe were selected from their respective groups, and 50 maps were simulated for both. From this point, the remaining metals were simulated by Co-DSS, ordered by their correlation coefficient with the previous element in the chain: Cr was simulated from As, Ni was simulated from Fe, Mn was simulated from Ni, and so on, until all the elements were simulated. The order in which the simulation was carried, starting from Fe, was: Fe, Ni, Mn, Pb, Cd, Zn, Co and Cu.

The second method, which is being used for the first time, also divided the variables into groups, based on their correlations with the PCs. Cd, Co, Cu, Fe, Mn, Ni, Pb and Zn were significantly correlated to the first principal component, PC1, while As and Cr were highly correlated with PC2. The first step was to model the spatial structure of the PCs, and simulate their own maps of values for the study area, using DSS. From these maps, the metal elements were simulated by Co-DSS. Each of the methods produced 50 maps of concentrations for the 10 heavy metals. These were then compared in terms of the reproducibility of the histogram and variogram of the samples, and later on, the maps of average and standard deviation values were compared.

### 3. Results and Discussion

The distributions of the concentration values of heavy metals on the biomonitors revealed histograms with positive skewness, with significant outlier values on the right tail of the distribution. For elements like As and Co (and Cr, to some extent), the skewness was not very noteworthy, while for the other elements, especially Mn and Ni, the histograms were quite skewed. This led to a significant difference between the average and median values of the distributions in the latter metals, as portrayed in Table 1. These statistics are measured in parts per million (ppm), except the coefficient of variation which is adimensional.

**Table 1 – Basic descriptive statistics on the metal concentrations measured in the biomonitors.**

	Average	Median	Standard Deviation	Coef. Variation	Skewness Coef.
<b>As</b>	4.16	3.36	2.87	0.69	1.55
<b>Cd</b>	0.27	0.22	0.26	0.95	3.35
<b>Co</b>	1.63	1.61	0.36	0.22	1.16
<b>Cr</b>	2.79	1.95	2.01	0.72	1.87
<b>Cu</b>	9.20	7.02	6.73	0.73	2.62
<b>Fe</b>	1206.87	834.77	1153.79	0.96	3.69
<b>Mn</b>	40.00	19.81	87.60	2.19	5.05
<b>Ni</b>	1.49	1.10	1.68	1.13	4.55
<b>Pb</b>	14.82	9.72	17.42	1.18	3.34
<b>Zn</b>	59.37	37.89	65.81	1.11	3.69

The average values were compared to other published studies, and it was concluded that the average levels of As in the region of Lisbon and south margin of the Tagus are similar or even higher than the levels of highly urbanized and/or industrialized areas in northern France (Cuny et al., 2004) and in the Turkish Aegean region (Yenisoy-Karakaş & Tuncel, 2004). The levels of Cr, Cu, Fe, Pb and Zn seem to be acceptable, as they are generally lower than the concentrations found in those two studies, and are comparable to those obtained for a region in center Italy (Brunialti & Frati, 2007). Cd and Ni showed to have quite low average values in Lisbon in comparison with those studies. Finally, it was found only one term of comparison for both Co and Mn, and the concentrations in Lisbon were significantly lower, so one might conclude that the average values are normal, although it would be better to have more references.

On the left of Figure 2 it can be seen a result of the PCA: the projection of the original variables (the concentrations of heavy metals) on the plane formed by F1 and F2, the two first Factorial Axes. The coordinates of the metals on this plane, which correspond to PC1 and PC2, show that two groups are formed, in terms of correlation coefficients: Cd, Co, Cu, Fe, Mn, Ni, Pb and Zn, which are projected very close to each other, and As and Cr, also close to one another. The two groups form an angle of approximately 90°, which means that they are independent.

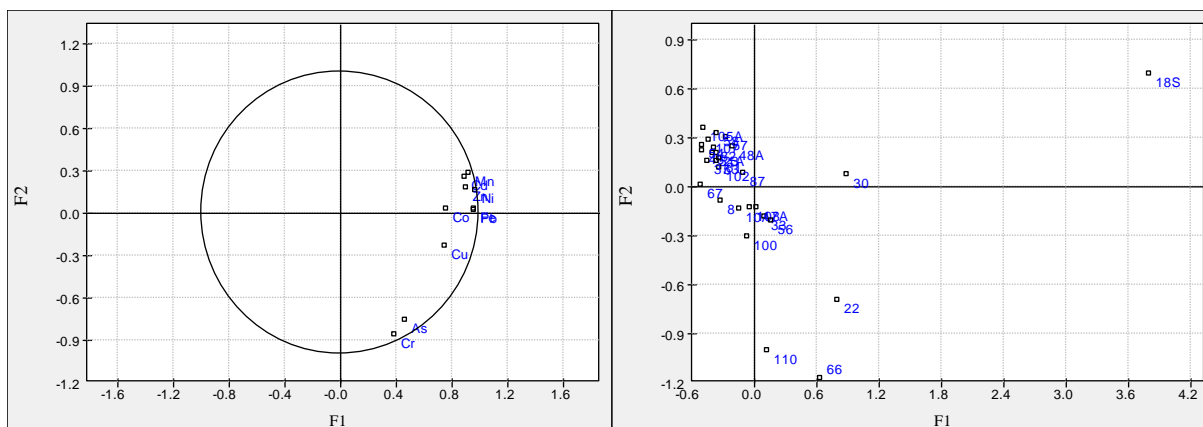


Figure 2 – Projection of the variables (left) and of the samples (right) on the 2 first Factorial Axes, F1 and F2.

Besides synthesizing the behavior of the 10 original variables, the two first axes also allow the identification of the samples with the highest concentrations in metals, which are also responsible for the construction of the main factorial axes of the PCA. On the right side of Figure 2 it can be seen that the sample with the reference 18S is the one with the largest contribution to F1, and, to a lesser extent, samples 30, 22 and 66. 18S was collected from the village of Paio Pires, Seixal, near the place where a steel mill used to operate, and it has the highest concentrations in most of the metals associated with F1, by far the most anomalous values in the histograms of these metals (with the exception on Co and Cu). Sample 30 was collected from a nearby place, the industrial area of Barreiro, while sample 22 was collected at Avenida da Liberdade, a large avenue in the center part of Lisbon, with intense automobile traffic. Sample 66 was collected at Vale Fundão, Chelas, in an urban park, and it not easy to say what is the origin of the high values of metal concentrations, and especially why it is not reaching sample 67, collected in the vicinity of 66, to the south. A possible source may be the area of Braço de Prata, which is relatively close to Vale Fundão (around 1 km), and where it was produced military equipment in past decades. Maybe the geographical attributes of the area or the dominant winds create special conditions that do not allow the metals in the atmosphere to reach the place where sample 67 was collected.

Regarding the metals associated with F2, it can be seen that samples 66, 110 and 22 are the ones with the highest concentrations. Besides being rich in the elements associated with F1, sample 22 also has high concentration on Cr, while 66 has high concentrations on both Cr and As. 110 was collected in Sobreda, Almada, near a farm, and is rich in As, so one might think that arsenical pesticides may be the cause of the high concentration.

Table 2 – Directions of highest (Dir 1) and lowest (Dir 2) spatial continuity, and respective variogram ranges (a1 and a2). The sill, C, and the nugget effect, C0, are also given.

	As	Cd	Co	Cr	Cu	Fe	Mn	Ni	Pb	Zn
<b>Dir 1</b>	-45°	45°	45°	-45°	45°	45°	45°	45°	45°	-45°
<b>a1</b>	24000	10800	9000	17000	10000	12000	8500	7500	10800	10000
<b>Dir 2</b>	45°	-45°	-45°	45°	-45°	-45°	-45°	-45°	-45°	45°
<b>a2</b>	12500	4000	7000	7500	5000	7500	4000	4500	5000	6000
<b>C</b>	7.97	0.06	0.13	3.88	43.68	1283679	7399.86	2.74	292.50	4176.61
<b>C0</b>	0	0.01	0.02	0	12.00	0	0	0	49.34	761.72

The calculation and modeling of the experimental variograms were carried out for all the variables, and the corresponding parameters of the fitted model can be found on Table 2. As indicated, 45° is the direction of highest spatial continuity for the variables Cd, Co, Cu, Fe, Mn, Ni and Pb, while for As, Cr and Zn, it is the direction -45°. The variograms of PC1 and PC2 were also modeled, because the simulation based on the PCs required it, and they were very similar to the variograms obtained for Fe and As, respectively.

The search for auxiliary variables was not successful, as no significant correlations between the metal concentrations and those variables were found. This does not mean that in reality the chosen variables are not associated with the emission of heavy metals to the atmosphere. The low correlations obtained were probably a result of the type and quality of data available. For example, the soil use map on which this study was carried was the CORINE Land Cover 2006 (European Environmental Agency, 2006), which does a classification of soil use from satellite images, and can only classify areas with an extension above 25 ha. This brings some limitations to the classification process, which sometimes puts very different soil use types into one category. Other data used were the energy consumption by municipality, which did not allow to distinguish between samples collected in the same municipalities. One final example is the distance to roads, which also did not bring any interesting results, probably because the samples were not collected at the best locations in terms of vicinity to the main roads of the study area (excluding sample 22). The fact that only 28 samples were available was probably also an important constraint to this approach, and as a consequence of these results, no auxiliary variables were used in the simulation process.

Simulation took place by two different approaches, as already mentioned. Next are some examples of simulations of the variable Ni, as simulated from Fe, Ni(Fe), and from PC1, Ni(PC1). For each approach, the 3 maps presented are different scenarios of concentrations distributions, with the same probability of occurrence. As seen on these figures, the maps in each group are significantly different from each other, which is an advantage of the simulation algorithm over the estimation (kriging) one.

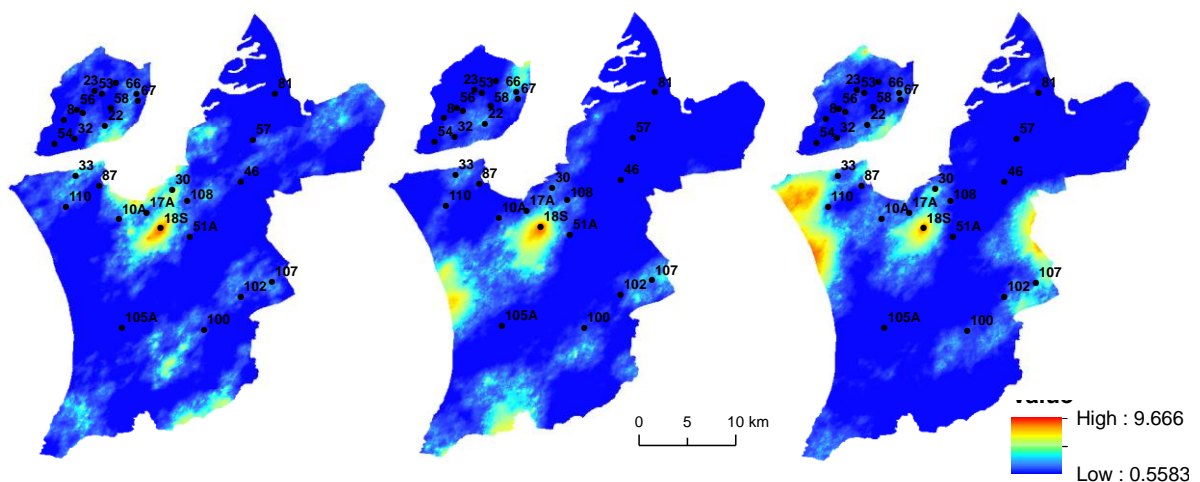


Figure 3 – Maps of 3 simulation of Ni, conditioned to Fe. The scale of concentration values is in ppm.

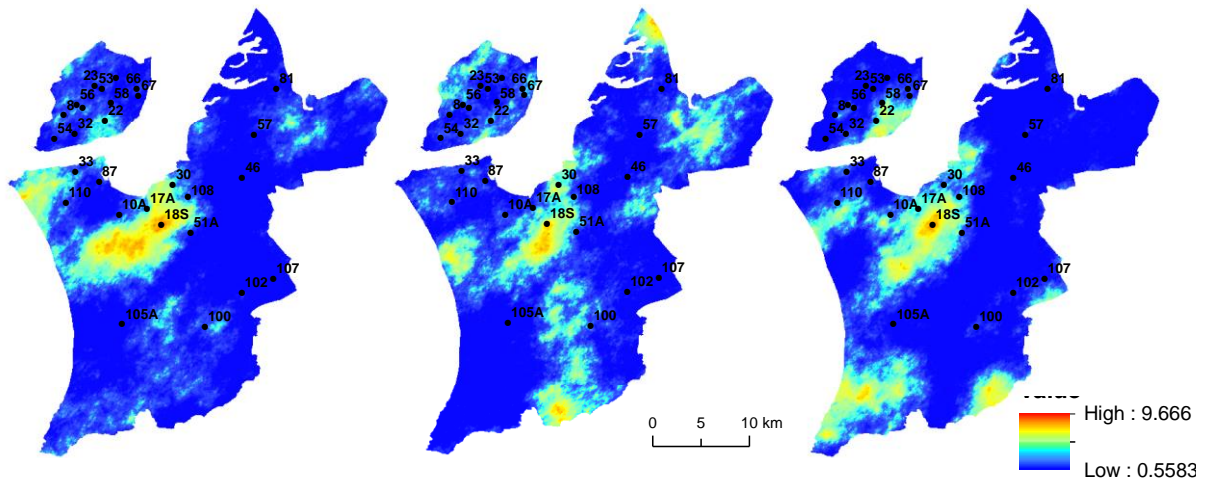


Figure 4 – Maps of 3 simulations of Fe, conditioned to PC1. The scale of concentration values is in ppm.

The comparison between each approach was done, at first, using the histograms and variograms of the simulations, comparing them to the histogram of the samples and the variogram model fitted to the experimental one. In general, the results obtained with both approaches were not significantly different, as can be seen on Figure 5 and 6. In both cases the variograms of the simulations grow under the model curve, which means that they have higher ranges (higher spatial continuity), and reach a lower sill than the model's one, which means that the simulations do not reproduce too well the variability of the samples.

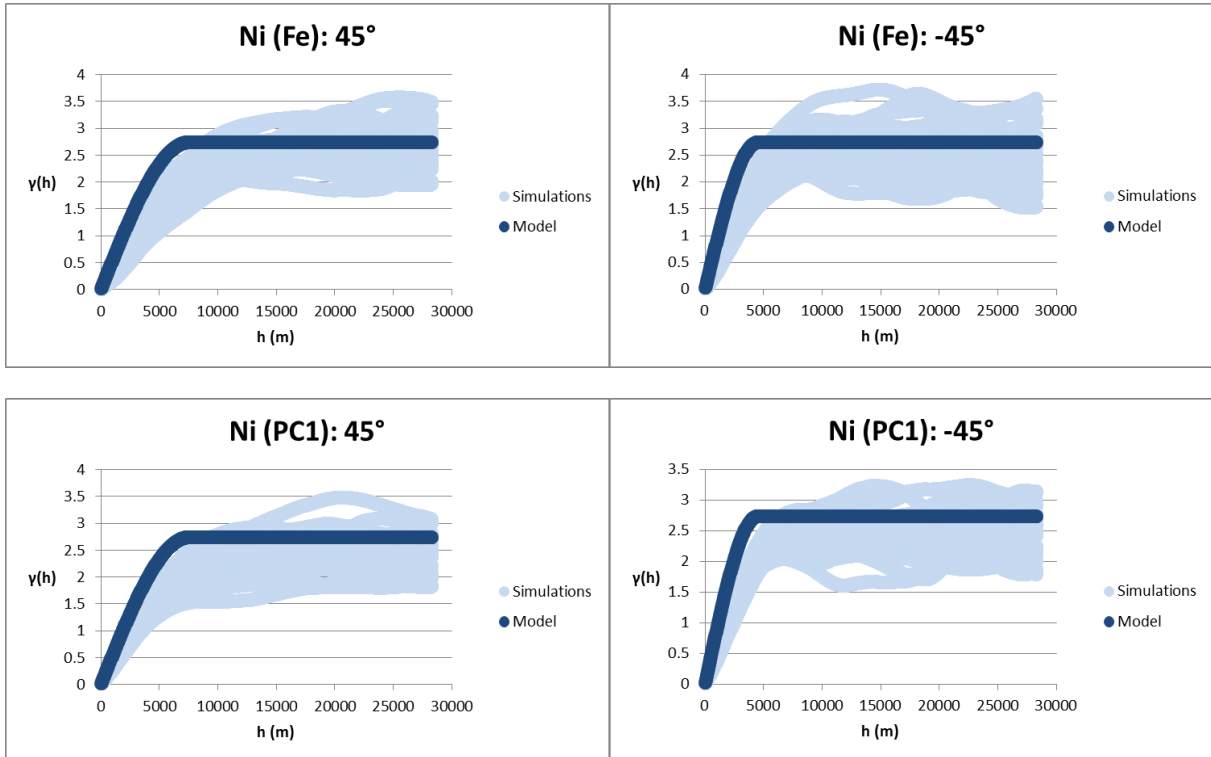


Figure 5 – Comparison between the fitted model of variogram and the variograms of 25 simulations.



Table 3 – Limits of the classes of the histograms of Ni, in ppm.

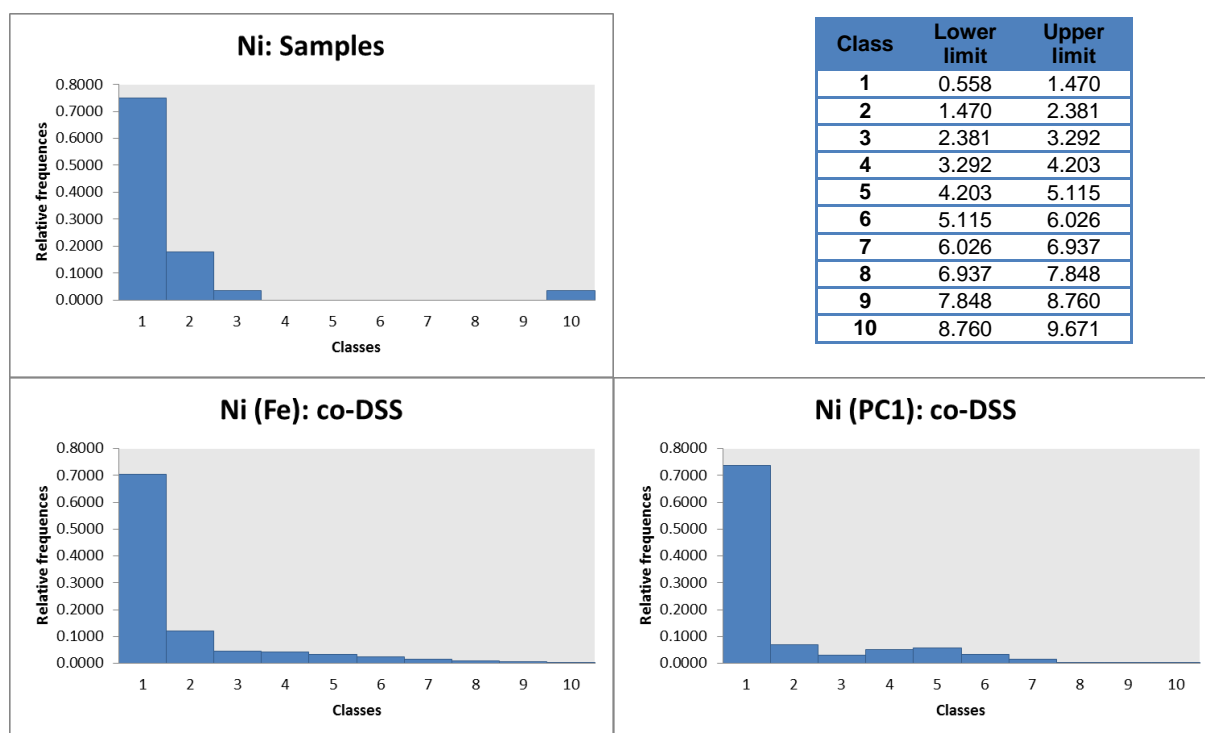


Figure 6 – Comparison between the samples histogram and the histograms of the two simulation approaches. The limits of the classes are presented in .

Looking at the histograms, it can be seen that the original one is highly skewed, with many empty classes. The algorithm could not reproduce those empty classes, which means that there was an attenuation of the variability, as also seen on the variograms. However, the form of the histograms of two simulations is identical to the form of the samples histogram. It can be noted both through the histograms and the simulated maps of Figure 3 and 4 that the middle classes of the concentration values are more densely populated in the simulation conditioned to PC1, as the light blue and green colors are more significant in the maps of Ni(PC1).

In general, the simulation of the variables associated with PC1 had problems similar to the simulation of Ni, namely the underestimation/overestimation of the variance (sill). Metals like Ni and Mn, with highly skewed histograms and high variability had an underestimation, while for the other metals there was an overestimation. This was probably due to the fact that Ni and Mn were simulated first, and influenced the following simulations. The same happened when simulating from PC1.

For As and Cr, which have more symmetrical histograms, generally there was not a problem with the reproduction of the sill. Both the variograms and histograms were well reproduced for the two simulation approaches, and no significant differences were found between the approaches.

For the second part of the comparison, maps of average and standard deviation values were computed. The higher average values for Ni are located, as expected, near samples 18S, 30, 22 and 66, which means that the simulation process took into consideration the values of the samples when simulating in their vicinity, which is something that is expected from the method. In this comparison, it was also observed a slight difference between the approaches, in the case of the group of 8 metals. In Figure 7 and 8 it can be seen that the map of average values for Ni(Fe) has wider dark blue areas, and the area of high values near 18S has more reddish colors. This means that Ni(Fe) has more



extreme average values (on both the lower and higher ends of the distribution) than Ni(PC1). In general, metals associated with PC1 revealed the same type of result, which means that when a PC synthesizes many variables, the simulation of those variables conditioned to the PC results in attenuation of the maps (and histograms). This was somewhat seen on the histogram of Ni(PC1), which had slightly more values in the middle classes, than Ni(Fe). In terms of the standard deviation maps, the lower uncertainties were found near the samples, because the algorithm simulates values with more confidence where it has more information available. It can be seen that the dark blue areas along with the orange/red ones are more prominent for Ni(Fe), which are the areas with the lowest and highest variability, respectively. In Ni(PC1), there are more light blue and greenish areas, which means that the variability is more homogeneous than in Ni(Fe).

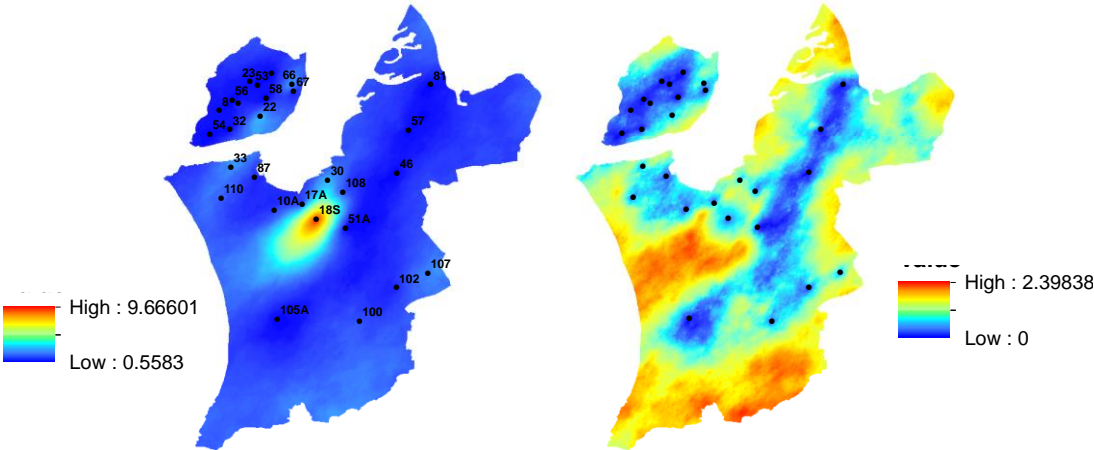


Figure 7 – Maps of average (left) and standard deviation values (right) for Ni(Fe), in ppm.

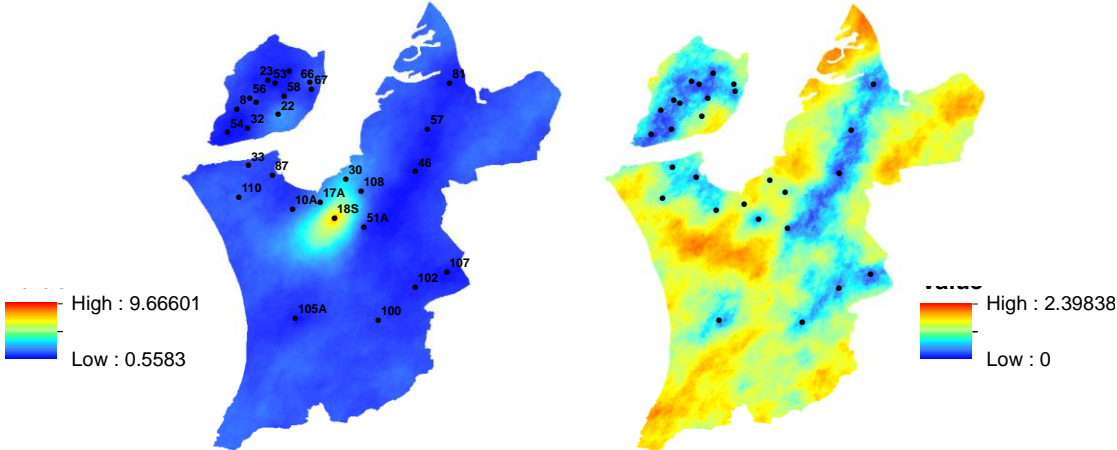


Figure 8 – Maps of average (left) and standard deviation values (right) for Ni(PC1), in ppm.

**4. Conclusions**

This work allowed to conclude that in general, the average concentration levels on atmospheric deposition in Lisbon and the south margin of the Tagus are acceptable for the studied metals, with the possible exception of As. The areas which were identified as being more relevant in terms of air pollution with the elements Cd, Co, Cu, Fe, Mn, Ni, Pb and Zn were the village of Paio Pires (Seixal),

the industrial area of Barreiro, Avenida da Liberdade and Vale Fundão (Chelas). For As and Cr, the main areas identified were Vale Fundão, Avenida da Liberdade and Sobreda (Almada). Different anthropogenic sources were identified as possible origins for the high concentration levels, but the search for auxiliary variables to help the process simulation returned no significant correlations between the chosen factors and the concentrations. This was probably due to the quality and quantity of the available data.

In terms of simulation, the two approaches produced satisfying results, with good overall variogram and histogram reproduction, with some limitations imposed mainly by the highly skewed concentrations distributions of the biomonitoring data. It was concluded that the simulation from the first PC resulted in slightly more attenuated maps and histograms, and with more homogeneous variability across the study area. This was probably due to the fact that PC1 synthesizes many variables (eight), and the combination of such variables leads to a more attenuated synthetic variable. However, the results of the simulation conditioned to the PCs were also good, and this new approach has the advantage of allowing the simulation of different variables without having to simulate a lot of other variables. It only requires that the PC is simulated.

## References

- Bargagli, R., Mikhailova, I., 2002. Accumulation of inorganic contaminants. In: Nimis, P. L., Scheidegger, C., Wolseley, P. (Eds.), *Monitoring with Lichens – Monitoring Lichens*. Kluwer/NATO Science Series, Dordrecht: 65-84.
- Brunialti, G., Frati, L., 2007. Biomonitoring of nine elements by the lichen *Xanthoria parietina* in Adriatic Italy: a retrospective study over a 7-year time span. *Science of the Total Environment* 387: 289–300.
- Cuny, D., Davranche, L., Thomas, P., Kempa, M., Van Haluwyn, C., 2004. Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry* 49: 391–401.
- Figueira, R., Sérgio, C., Sousa, A., 2002. Distribution of trace metals in moss biomonitors and assessment of contamination sources in Portugal. *Environmental Pollution* 118: 153–163.
- Franco, C., Soares, A., Delgado, J., 2006. Geostatistical modelling of heavy metal contamination in the topsoil of Guadiamar river margins (S Spain) using a stochastic simulation technique. *Geoderma*, 136: 852–864.
- Järup, L., 2003. Hazards of heavy metal contamination. *British Medical Bulletin* 68: 167-182.
- Reis, A. P., Patinha, C., Ferreira da Silva, E., Sousa, A. J., Figueira, R., Sérgio, C., Novais, V., 2010. Assessment of human exposure to environmental heavy metals in soils and bryophytes of the central region of Portugal. *International Journal of Environmental Health Research*, 20(2): 87–113.
- Soares, A., 2001. Direct sequential simulation and cosimulation. *Mathematical Geology* 33: 911–926.
- Yenisoy-Karakaş, S., & Tuncel, S. G., 2004. Geographic patterns of elemental deposition in the Aegean region of Turkey indicated by the lichen, *Xanthoria parietina* (L.) Th. Fr. *Science of the Total Environment*, 329: 43–60.