OPTIMAL REGIONAL SAMPLING NETWORK TO ANALYSE ENVIRONMENTAL POLLUTION BY HEAVY METALS USING INDIRECT METHODS. CASE STUDY: GALICIA (NW OF SPAIN)

CARME HERVADA-SALA[1], EUSEBI JARAUTA-BRAGULAT[2]

[1] *Dept. Physics and Nuclear Engineering. EUETIT (UPC), Colom 1, 08222-Terrassa.*
Carme.Hervada@upc.es

[2] *Dept. Applied Mathematics III. ETSECCPB (UPC). Jordi Girona 1-3, 08034-Barcelona.*
Eusebi.Jarauta@upc.es

Website: `ftp://ftp-urgell.upc.es/Matematica/E.Jarauta/Geoenv2002`

ABSTRACT

Environmental pollution by heavy metals is a red-hot issue. It is being studied from many points of view, as it is not only an environmental problem but also a public health matter. The effect of pollution by heavy metals can be assessed directly, that is measuring heavy metal concentration in soils, or using indirect methods, that is measuring heavy metal contents on living beings of regional ecosystem, in particular on plants. One of the organisms that have proved to be the most faithfully and useful to do so are moss. So, heavy metal environmental pollution can be studied by taking moss samples and measuring their heavy metal contents.

The aim of this work is to show the use of geostatistical tools in environmental pollution analysis applied to a case study of environmental pollution by heavy metals in Galicia (north west of Spain). To do so, two different information in that zone are available: on one hand, measures of heavy metal concentration in moss (Scleropodium purum), whose location points are known, also their level. On the other hand, situation of polluting sites (industrial areas and towns) and their classification taking into account their polluting capacity. This information allows assessing not only for the regional pollution, but also for its scattering. From this and using geostatistical tools, sampling network is being improved. Data set consists of 71 sample points where concentration of ten elements (Al, As, Co, Cr, Cu, Fe, Hg, Ni, Pb and Zn) is measured. For each of them classical statistical analysis is done. Furthermore, spatial variability is studied using a new methodology based on Fast Fourier Transform (FFT), which allows finding covariance matrix using all variables at the same time. FFT methodology improves the classical and tedious geostatistical methodology based on variogram and cross-variogram modelling to find data spatial variability. Finally contour maps of environmental pollution by heavy metals in Galicia are presented.

1. Introduction and objectives

Galicia is a region located at the northwest of Spain. It is about 29434 km^2 large. It ranges, approximately, between 7° and 9° western meridians and 40° and 42° northern parallels. The outline of this region is gently undulated, with hills and valleys; this smoothness defines its landscape with a series of high and low regions at several levels. So Galicia's landscape is full of high and low areas. The highest areas are in its east border.

Forests, however original man has regularly modified vegetation, cover the most part of Galicia. The most widespread trees are oak, chestnut, birch, cork and ilea. From some time on there have been also planted pine and eucalyptus. Another interesting aspect in Galicia is that their towns are small and scattered all over the whole country. Recently some of the towns have grown due to the enlargement of some industrial zones. The main industrial activities are cars (located at the north), woodwork, textiles, and craftsmanship.

As it is well known, the increasing of industrial activities implies a pollutant impact on the environment. Nevertheless, our society demands a quality of life compatible with technical progress, without renouncing to it. One of the most important contributions the environment pollution caused by industry is the presence of heavy metals in the air, which fall down when it rains and then are incorporated by leaving beings. In this article the presence of heavy metals from industrial origin are studied, in fact samples of them are measured on some moss: *Scleropodium purum* (Hedw.) Limpr. The metals that are measured are aluminium, cobalt, chromium, copper, iron, mercury, nickel, lead and zinc, and also the metalloid arsenic.

Accumulation of heavy metals over large areas and long periods causes damage to living organisms and it must be carefully controlled; it is also important to know the effects of these contaminants. To assess the pollution caused by metals there are two different methods: the direct one, which consists of measuring their concentration in the air or in soil, and the indirect one, which consists of studying their presence in some living beings. If previous monitoring is correctly done, indirect method can be useful in environmental assessment, because it is easiest and cheapest. This monitoring has been done in Galicia with moss (see reference [2]).

The aim of this work is to build up contouring level maps of pollution by heavy metals using geostatistical methods; to do so, we take into account measures in several moss samples. The method used is kriging on a regular grid with correlogram tables obtained by applying the Fast Fourier Transform (FFT) methodology (see reference [5]).

2. Data set: description and analysis

Full database can be found in our website. Professor J.A. Fernández, from Universidad de Santiago de Compostela, has supplied data. Data consists of 71 samples of all ten (see above) elements' concentration (ppm). The number of sampling sites was the equivalent to a density of 2.6-samples/1000 km^2, higher than the density recommended for such studies at a regional scale. Sampling was carried out in 1995 (April-July) and covered almost all Galicia, with a higher density at most industrial areas. The concentration of Al, Co, Cr, Cu, Fe, Ni, Pb and Zn in moss extracts were determined using flame absorption spectrophotometry and Hg and As were determined using atomic fluorescence. Figure 2.1 shows a scatter plot of the location points; the co-ordinates are UTM scaled and in kilometres. Sampling points were taken at different levels, between 72 m and 1014.5 m high. In the light of sampling points and Galicia's dimension, a kriging grid of 9x9 nodes has been built. Distances between nodes are 21x23 km.

To be able to carry out a bidimensional geostatistical study, we tried to find out a possible functional relation between data and altitude. The conclusion is that this dependence does not exist. Figure 2.2 shows, as an example of that, the scatterplots for Al and Cu concentration versus altitude (ALT); this figure shows also the regression line, which is quite horizontal. The corresponding hypothesis test shows that it is not possible to reject the fact that correlation between data and altitude does not exist.

Figure 2.1.- Scatterplot of sample locations.

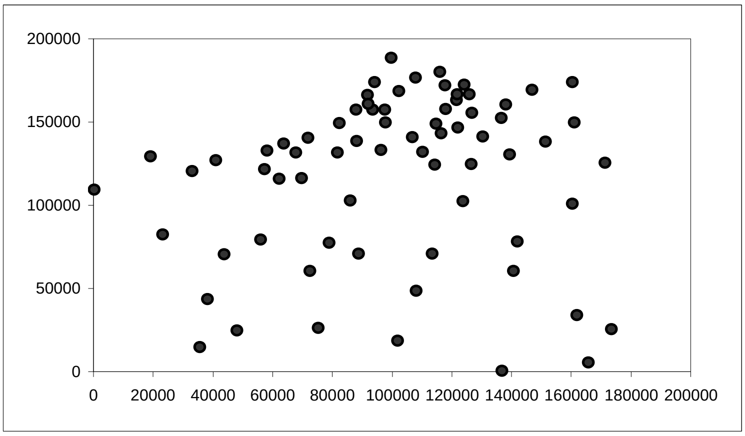


Figure 2.2.- Scatterplots: Al (a) and Cu (b) versus altitude (ALT).

(a) (b)

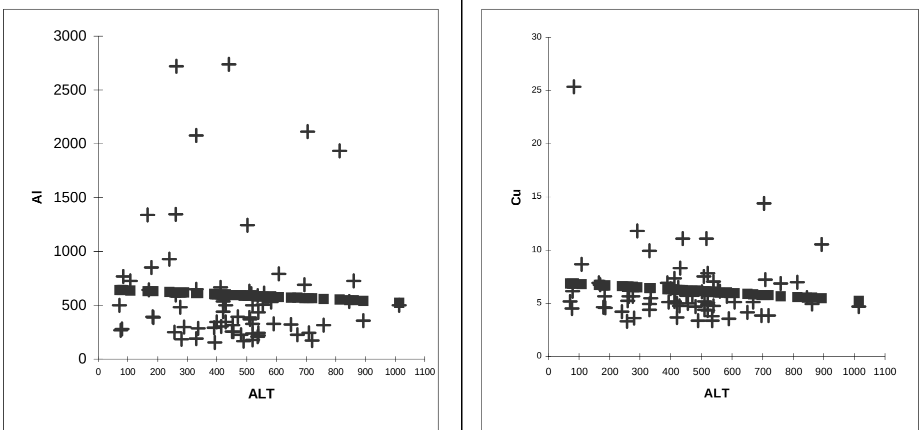


Table 2.1 shows the average concentration distributed at different levels, at altitude intervals of 100 m; the ALT values are the averages in the corresponding interval. In table 2.2, there is a statistical descriptive analysis of data. In table 2.3, correlation coefficients of the ten elements and altitude are shown. Finally, in figure 2.3 there are shown the variable histograms (element concentration). For additional information about this data set, see reference [2].

Table 2.1.- Concentration averages (ppm) at different altitude levels (m) for each element.

ALT	Al	As	Co	Cr	Cu	Fe	Hg	Ni	Pb	Zn
77.8	454.4	0.298	0.422	1.190	10.287	175.0	0.076	2.178	3.26	59.7
165.4	723.8	0.447	0.709	1.545	6.202	769.0	0.074	2.313	7.90	54.3
265.6	850.6	0.317	0.616	1.605	5.740	580.4	0.028	1.775	5.23	57.3
358.4	571.3	0.166	0.500	1.663	6.177	502.2	0.039	1.981	3.25	48.1
440.3	553.7	0.230	0.411	1.455	5.870	539.8	0.034	1.948	7.87	59.9
533.7	453.4	0.272	0.297	1.171	5.740	459.2	0.037	1.602	4.20	61.1
655.9	508.2	0.260	0.442	1.216	4.563	398.4	0.039	1.489	4.71	52.0
723.5	712.1	0.148	0.169	1.232	8.098	1218.6	0.033	1.610	14.54	69.7
885.2	810.6	0.432	0.298	1.133	6.547	435.4	0.045	1.629	2.74	64.1

Table 2.2.- Descriptive statistics of sample data set (ppm).

	Al	As	Co	Cr	Cu	Fe	Hg	Ni	Pb	Zn
Mean	596.6	0.277	0.419	1.364	6.229	541.9	0.041	1.820	5.766	58.66
Median	436.5	0.204	0.307	1.105	5.553	388.4	0.034	1.722	3.658	57.27
Minimum	156.6	0.005	0.060	0.176	3.316	139.8	0.002	0.539	0.04	31.22
Maximum	2740.7	1.276	2.498	4.769	25.352	4028.9	0.203	5.018	60.42	117.67
Standard Dev.	548.3	0.263	0.413	0.833	3.122	580.9	0.030	0.845	8.637	18.14

Table 2.3.- Correlation coefficients of elements concentration and altitude (ALT).

	ALT	Al	As	Co	Cr	Cu	Fe	Hg	Ni	Pb	Zn
ALT	1.000										
Al	-0.047	1.000									
As	-0.049	0.347	1.000								
Co	-0.290	0.696	0.283	1.000							
Cr	-0.174	0.634	0.185	0.605	1.000						
Cu	-0.116	0.337	0.036	0.254	0.200	1.000					
Fe	0.021	0.768	0.219	0.511	0.521	0.381	1.000				
Hg	-0.217	0.243	0.233	0.414	0.209	0.107	0.175	1.000			
Ni	-0.267	0.525	0.253	0.643	0.395	0.662	0.400	0.208	1.000		
Pb	-0.002	0.127	-0.087	0.035	0.135	0.213	0.361	-0.083	0.198	1.000	
Zn	0.139	0.062	0.063	-0.061	-0.226	0.264	-0.003	0.061	0.176	0.124	1.000

Figure 2.3.- Histograms of elements' concentration (ppm): (a) Al, (b) As, (c) Co, (d) Cr, (e) Cu, (f) Fe, (g) Hg, (h) Ni, (i) Pb, (j) Zn.

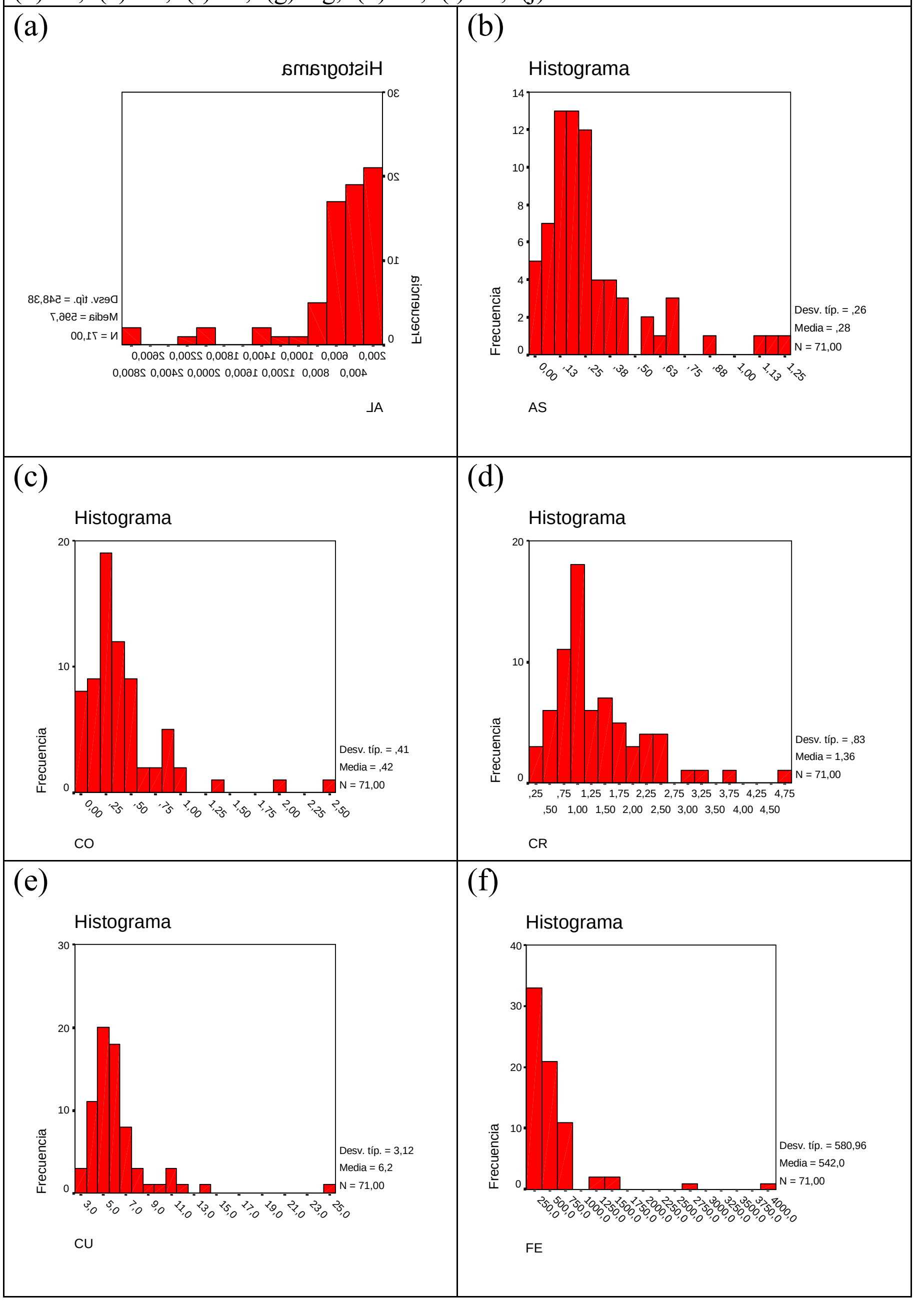
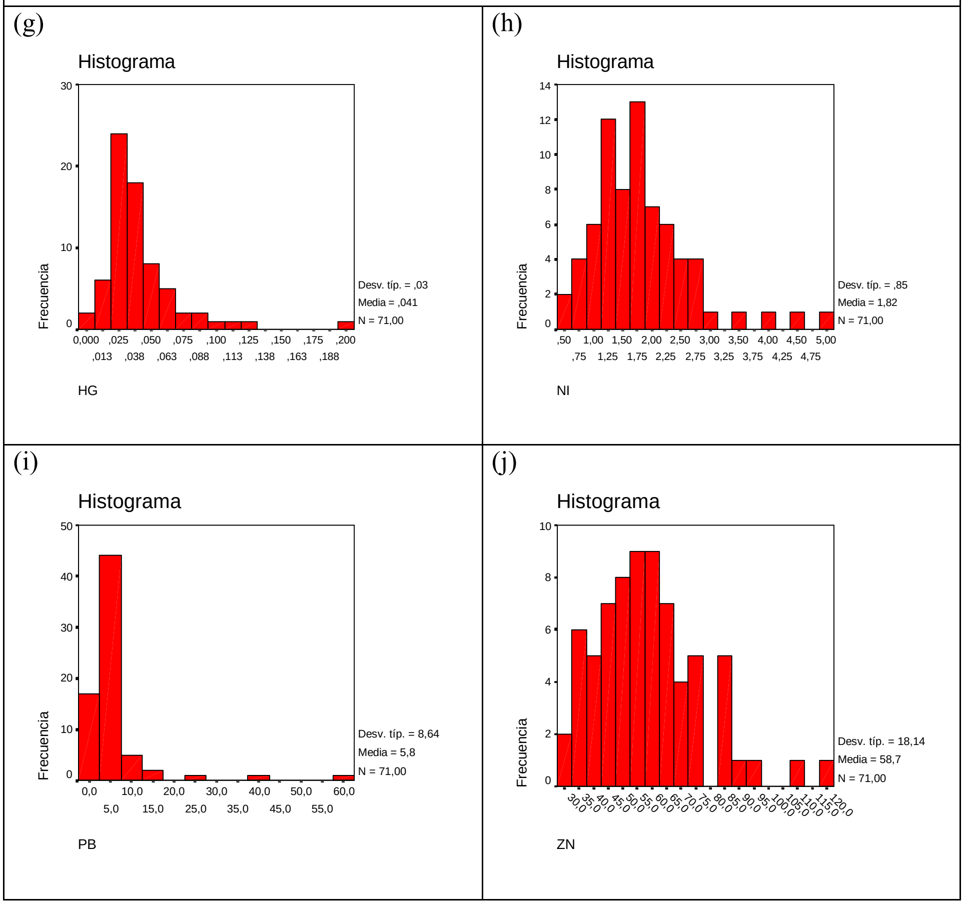


Figure 2.3 (Cont.)



3. Geostatistical analysis: pollution maps

The first step in the geostatistical study, which is the most important goal of this work, is to calculate the data Normal Score Transform (NSCT), according to the GSLIB procedure (see reference [1]). In this transformation, as Al and Fe do not have a Gaussian cumulative distribution function, some adjustments of their ties have had to be done. Parameters used in this program are shown in table 3.1.

The second step, which is the equivalent to calculate and model variograms in classical geostatistics, is the building of the initial correlogram matrix using Fast Fourier Transform (FFT) following the methodology established by Yao and Journel (1998) and Ma and Yao (2001); see references [5] and [4]. This matrix consists of a 10×10-

block matrix, which has in its diagonal the auto-correlations and the remaining the cross-correlations. So, we obtain 55 different correlation maps. The number of grid points in each map has to be of $1+2^n$; in this case n = 5, that is, we have a 33×33 element matrix. The correlations have been interpolated using a size 10 smooth window and then multismoothing using all correlations and variables with size 3 maximum half window has been carried out to have the final correlation. Some of those 55 maps are hanged at our website. Corresponding parameters are shown in tables 3.2, 3.3 and 3.4.

The third step is to krige on a regular grid using those correlation maps. Kriging has been done using the program KB2D modified by Hervada-Sala and Jarauta-Bragulat (2001) see reference [3]. After kriging, coordinates must be added taking in mind grid parameters; they are shown in tables 3.5 and 3.6. At last, back transformations of all results must be computed to recover original space and units. Parameters for that back transformation are shown in table 3.7. Figure 3.1 shows the contour maps obtained from the kriging grid with the right back transformed values.

Table 3.1: Parameters for NSCORE

```
molsagal.dat      \file with data
13  0             \columns for variable and weight
-900   900        \trimming limits
0                 \1=transform according to specified ref. dist.
hist1.out         \file with reference dist.
1   0             \columns for variable and weight
nsgal10.dat       \file for output
nsgal10.trn       \file for output transformation table
```

Table 3.2: Parameters for CORRMAP

```
molsagal.dat                \file with data
10 4 5 6 7 8 9 10 11 12 13 \number of variables: column numbers
-999 999                    \trimming limits
0                           \1=regular grid, 0=scattered values
33   33                     \if =1: nx,     ny
1 1                         \xsiz, ysiz
1 2                         \if =0: columns for x,y coordinates
corgal                      \file for correlogram output
16 16                       \nxlagl, nylag
5250 5750                   \dxlag, dylag
1 1                         \xtol, ytol (in the grid unit)
1                              \minimum number of pairs
```

Table 3.3: Parameters for INTPMAP

```
10                    -number of variables
33 33                 -num. of nodes in x and y directions
corgal                -file with sample corr
intpgal               -output file with interpolated correlogram
indbg                 -debug file
10                    -smooth window
0.1 0.01              -ratio of the inner and outer radius of fan
4
```

Table 3.4: Parameters for MULTSMTH

```
10                  -number of coregionalized variables
33 33               -number of nodes in x and y dir.
intpgal             -input file with original corr.map
mapagal             -output file of permissible corr.map
3                   -maximum half smoothing window size
0                    -minimum number of data for smooth.
```

Table 3.5: Parameters for KB2D

```
molsagal.dat      \file with data
1   2   4         \columns for X, Y, and variable
-999 999          \trimming limits
3                 \debugging level: 0,1,2,3
kb2d.dbg          \file for debugging output
krigal01.out      \file for kriged output
9 300   21000     \nx,xmn,xsiz
9 550  23000      \ny,ymn,ysiz
1 1               \x and y block discretization
1    8            \min and max data for kriging
2.11e4            \maximum search radius
1    2.302        \0=SK, 1=OK,  (mean if SK)
mapagal.1         \cov file
31 31
```

Table 3.6: Parameters for ADDCOORD

```
krigal10.out       \file with data
krigal10.dat       \file for output
1                  \realization number
9 300    21000     \nx,xmn,xsiz
9  550   23000     \ny,ymn,ysiz
1      1     0     \nz,...
```

Table 3.7: Parameters for BACKTRANS

```
krigal10.dat     \file with data
4                \column with Gaussian variable
-900 900         \trimming limits
bacgal10.out     \file for output
nsgal10.trn      \file with input transformation table
31.20 117.7      \minimum and maximum data value
1    0.05        \lower tail option and parameter
1  2             \upper tail option and parameter
```

Figure 3.1.- Contour maps of kriging results: (a) Al, (b) As, (c) Co, (d) Cr, (e) Cu, (f) Fe, (g) Hg, (h) Ni, (i) Pb, (j) Zn.

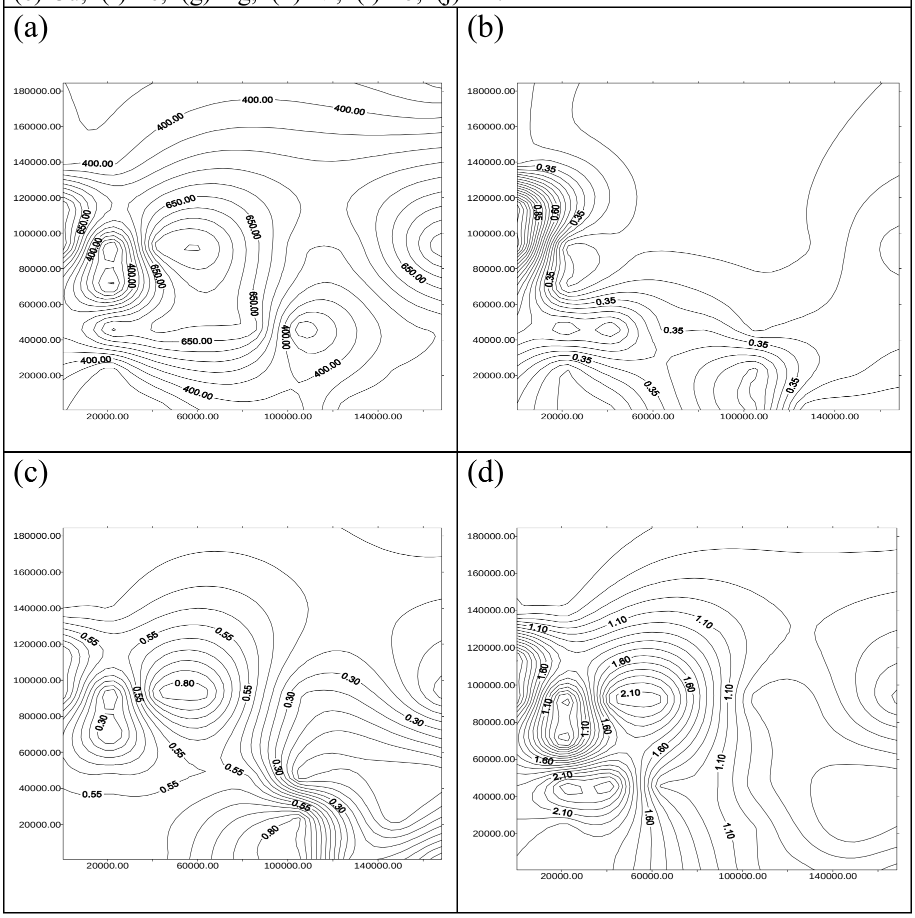
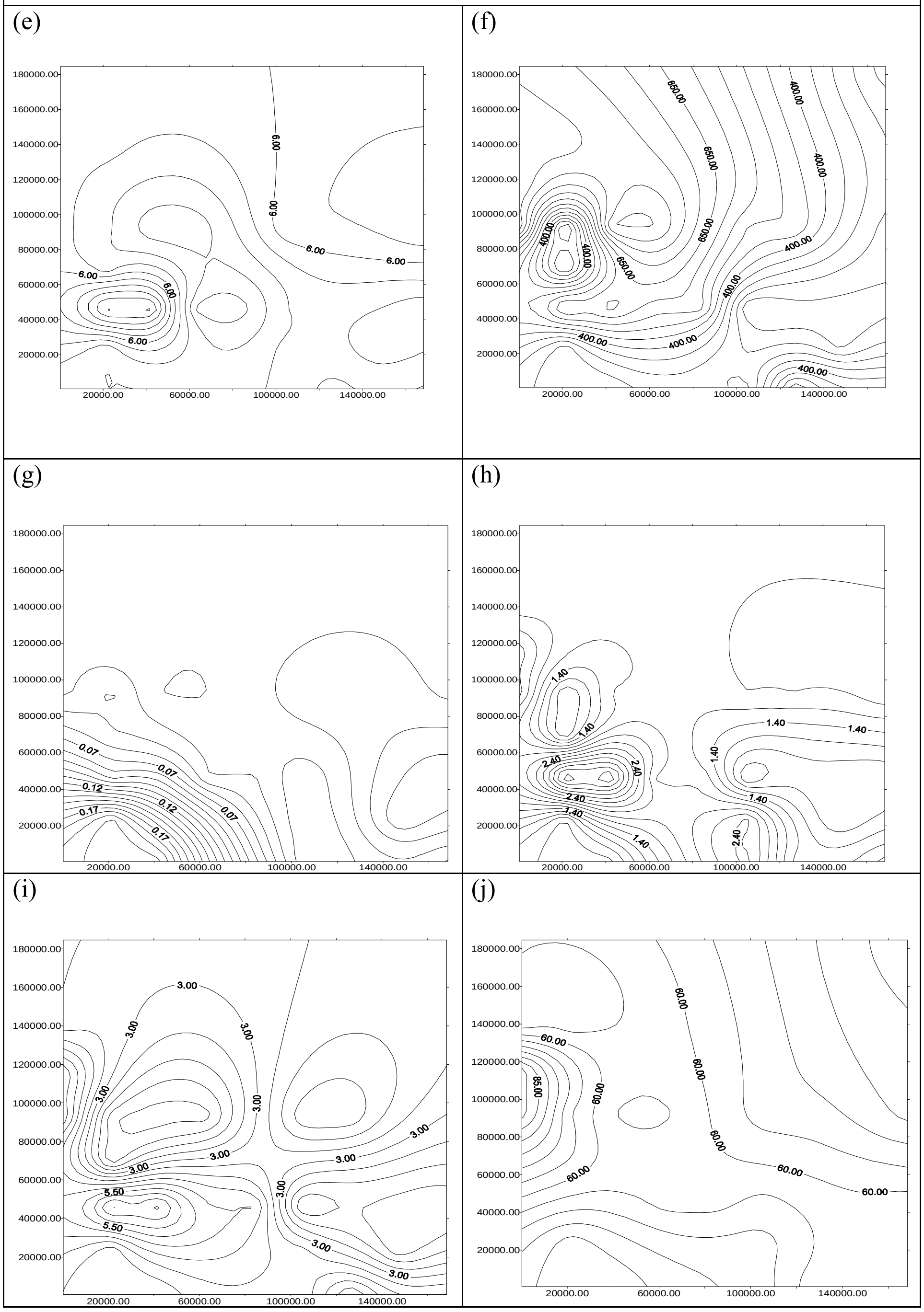


Figure 3.1 (Cont.)



4. Conclusions

The main conclusions of this work are the following:

1) It is possible to improve the statistical analysis of environmental pollution by heavy metals in Galicia done in [1], using a two-dimensional geostatistical analysis.
2) Sample density used in this study in not enough to reflect variability of environmental pollution, due to geography of Galicia; it is not possible to employ parameters fitted for a regional scale in that case.
3) Great problems arise with the use of classical geostatistical tools, based on variograms and cross variograms modeling. However, the use of modern FFT techniques allows for finding the full correlogram maps and so it is possible to krige adequately on a regular grid.
4) Finally, it has been possible to build contouring maps of all variables that reflect quite adequately the distribution and concentration of heavy metal pollution. This allows the design of a better sampling grid to control more accurately heavy metal pollution in that region.

5. References

[1] Deutsch, C.V. and A.G. Journel (1998). *Geostatistical software library and user's guide – GSLIB*. Oxford University Press, 1 CD + 369 pp.

[2] Fernández, J.A., A. Rey and A. Carballeira (2000). *An extended study of heavy metal deposition in Galicia (NW Spain) based on moss analysis.* The Science of Total Environment, vol. 254, pp 31-44.

[3] Hervada-Sala, C. and E. Jarauta-Bragulat (2001). Modifications to kb2d program in GSLIB to allow use of tabulated covariances calculated with Fast Fourier Transform method. Computers & Geosciences, vol.27, num. 07, pp 887-889.

[4] Ma, X and Yao,T. (2001). A program for 2D modeling (cross)correlogram tables using Fast Fourier Transform. Computers & Geosciences, vol.27, num. 07, pp 763-774.

[5] Yao, T. and A.G. Journel (1998). Automatic modelling of (cross)covariance tables using fast Fourier transform. Mathematical Geology, 30(6), 589–615.

Acknowledgements

Authors want to thank specially to professor J.A. Fernández, from the University of Santiago de Compostela (Spain), for furnishing the data set.