4th Imeko TC19 Symposium on Environmental Instrumentation and Measurements
Protecting Environment, Climate Changes and Pollution Control
June 3-4, 2013, Lecce, Italy

SAMPLING OPTIMIZATION FOR MONITORING CONTAMINATED SOILED

V.Pelillo[1], L. Piper[2], A. Lay-Ekuakille[1], G. Griffo[1], A. Lanzolla[3], G. Andria[3]

[1] Dipartimento d'Ingegneria dell'Innovazione, University of Salento, Via Monteroni, 73100, Lecce, Italy
[2] Dipartimento di Scienze Economiche e Matematico Statistiche, University of Salento, Via Monteroni, 73100, Lecce, Italy
[3] DEE, Polytechnic of Bari, Via Orabona 4, 70100 Bari, Italy
[vincenza.pelillo,luigi.piper,aime.lay.ekuakille,Giuseppe.griffo]@unisalento.it, lanzolla,andria@misure.poliba.it
http://smaasis-misure.unile.it, Ph.+39.0832.297822 Fax +39.0832.297822
Department of Innovation Engineering, University of Salento, Lecce, Italy, vincenza.pelillo@unisalento.it

Abstract: Atmospheric deposition on soil is a matter of continuous study. Deposited pollution can be determined in various ways. But the first step is sampling; in case of sample destruction or lost or unvailability, it could be necessary to overcome the issue by a specific prediction. The paper presents prediction technique of dioxin deposition on ground using a geostatistic method called "Kriging". Kriging code is a regression method used in the context of the space analysis (geostatistics) which allows to interpolate a quantity in the space, minimizing the mean square error. In the statistics context it is better well known as gaussian process. Knowing the value of a quantity in a few points in the space (for example the analyte concentrations taken in every town of an area), we can determine the quantity value in other points for which measurements do not exist. The performed algorithm displays better results than many other ones because of geostatistic properties adopted in the procedure. Analytical data are taken from a campaign carried out in Lecce province.

Keywords: soil contamination, sampling modelling, statistical parameters, uncertainty, geostatistics.

1. INTRODUCTION

Geostatistics is a discipline which promotes a method able to supply the best estimate of a quantity, in a prearranged domain, based on a preliminary analysis of the space variability of the experimental data belonging to the domain. Various types of approaches at the study of geostatistics exist:

- according to a first approach it is indispensable to start from the theory of the random functions to see the basic geostatistics estimate algorithm (Universal Kriging);
- according to another approach, instead, Kriging can be defined and implemented without any sign to its probabilistic features.

Independently of the followed approach, geostatistics agrees:

- to analyse the space structure of a quantity inside a prearranged domain on the basis of available experimental data;
- to adopt an appropriate model (probabilistic) representative of the characters of variability or space continuity of the quantity inside the fixed domain;
- to make the systematic estimate of quantity, considering the redundancies of the available data;
- to evaluate the reliability of the obtained estimates.

The technique treated in this work is Ordinary Kriging, which is a technique associated often with the B.L.U.E. acronym ("Best Linear Unbiased Estimator"). In other words the Kriging is a geostatistics estimator who has two important features:

- it is exact, i.e. returns, in the points where we have the space information, its true value;
- for every estimated value it provides the estimate variance, a parameter for the evaluation of the accuracy of the proposed estimates.

In Kriging, we use a model of probabilities in which the deviation and the error. Kriging [1] is an anticipatory method which is different from the other one because it produces a set of estimates with minimum error variance. The error variance, σ^2_R, of a set of k estimate can be written as:

$$\sigma^2{}_R = \frac{1}{k}\sum_{i=1}^{k}[\hat{v}_i - v_i - \frac{1}{k}\sum_{i=1}^{k}(\hat{v}_i - v_i)]^2 \qquad (1)$$

where v_i are the real values and $\hat{v}_i$ the esteemed values. Assuming an average error equal to zero, the previous equation becomes:

$$\sigma^2{}_R = \frac{1}{k}\sum_{i=1}^{k}(r_i - 0)^2 = \frac{1}{k}\sum_{i=1}^{k}[\hat{v}_i - v_i]^2 \qquad (2)$$

To minimize the variance of the error [2], it is necessary to introduce others parameters, adding the Lagrange parameter: Calculating the partial derivative and placing equal to zero we obtain the set of weights which minimize the error variance [3] [4] [5].

$$\tilde{\sigma}^2{}_R = \tilde{\sigma}^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \tilde{C}_{ij} - 2\sum_{i=1}^{n} wi\tilde{C}_{io} + + \underbrace{2\mu(\sum_{i=1}^{n} w_i - 1)}_{0} \quad (3)$$

This equation system, often called Ordinary Kriging System, can be written in matrix notation as:

$$\begin{bmatrix} \tilde{C}_{11} & ..... & \tilde{C}_{1n} & 1 \\ ..... & ..... & ..... & ..... \\ \tilde{C}_{n1} & ..... & \tilde{C}_{nn} & 1 \\ 1 & ..... & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ .... \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{C}_{10} \\ .... \\ \tilde{C}_{n0} \\ 1 \end{bmatrix} \quad (4)$$

To obtain the weights, both members of this last equation are multiplied for the inverse one of the matrix of covariance, C^1, obtaining:

$$W = C^{-1} \cdot D \quad (5)$$

Determined the model weights, the error variance value is obtained from the following equation:

$$\tilde{\sigma}^2{}_R = \sigma^2 - (\sum_{i=1}^{n} w_i \tilde{C}_{io} + \mu) \quad (6)$$

or, in matrix terms:

$$\tilde{\sigma}^2{}_R = \tilde{\sigma}^2 - W \cdot D \quad (7)$$

This minimized error variance is usually called as Ordinary Kriging Variance. The estimate of a quantity through Kriging method depends, however, on the covariance model Under the practical point of view, the model chosen for the random function is bound to the space continuity of the sampled data set. To be sure of the solution uniqueness, the system must be defined positive. Below he a qualitative graphic representation of the used variogram models.

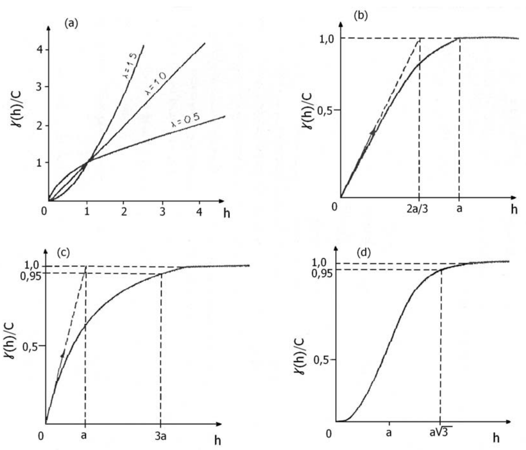

Fig. 1. Variogram Model: a. Power Model; b. Spherical Model; c. Exponential Model; d. Gaussian Model.

2. SAMPLING AND CHARACTERIZATION

Pollutant deposition on soil is an important issue that connect air pollution with soil. Pollutant measurement prediction is a topic of great interest in the area of environmental measurements and health protection. Stationary description of pollutants is an approach in which the amount of pollutant per time unit is considered constant. But in many circumstances, the amount of pollutants is subject to a flow delay that is not taken into account because of complexity. The general scheme is highlighted in fig.2 where we can see different phenomena. This kind of modeling can better reflect forced and spontaneous emissions of pollutants from industrial plants and natural processes respectively [6] [7].

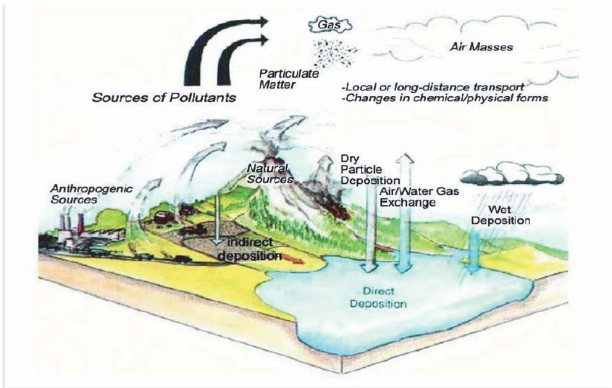

Figure 2. General scheme of pollutant deposition

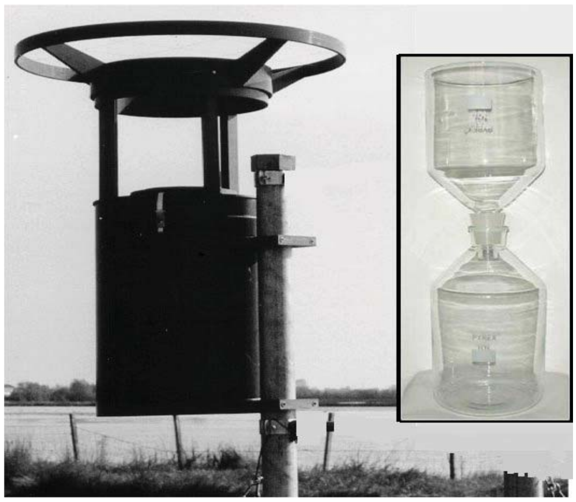

Figure 3. Deposimeter for sampling and its tube

Different deposimeters (fig.3) have been used to sample the concentrations of diverse pollutants, namely, dioxins, furans, aromatic hydrocarbons. Deposimeters are located in different areas of the Province of Lecce (Italy) as indicated in fig.4.

Figure 4. Deposimeter for sampling and its tube

3. PROPOSED GEOSTATISTIC APPROACH

It is important to notice that the correctness of obtained information by Kriging method depends strongly on the chosen variogram model [8] and on the isotropic features of the observed phenomenon. In the building of the experimental variograms, a different behaviour of the data depending on the direction of chosen investigation is very often discovered. This behaviour shows the presence of anisotropies which must be considered in the building of Kriging.

Isotropic model

The first approach to be followed in the building of an Experimental Variogram, when we haven't information about the space sample features, is the isotropic one.

The chosen Variogram model is:

$$\gamma(h) = c\left[1 - \exp\left(-\frac{3h}{a}\right)\right] \tag{8}$$

where *c* and *a* respectively represent the sill and the range. The determination of these two parameters has been led trying to have the minor difference between the experimental data and the model of variogram. In the following figure we have the results of calculation.

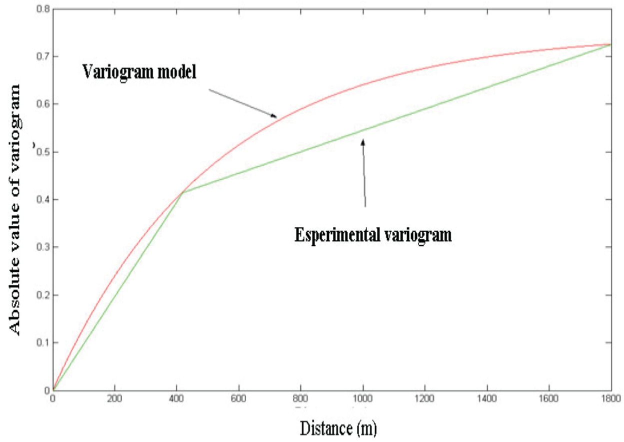

Figure 4. Representation of the exponential isotropic model of Variogram.

Anisotropic model

In this case, it is possible to proceed in the experimental determination of variogram from which we obtain the variogram model for the description of the relationships between data.

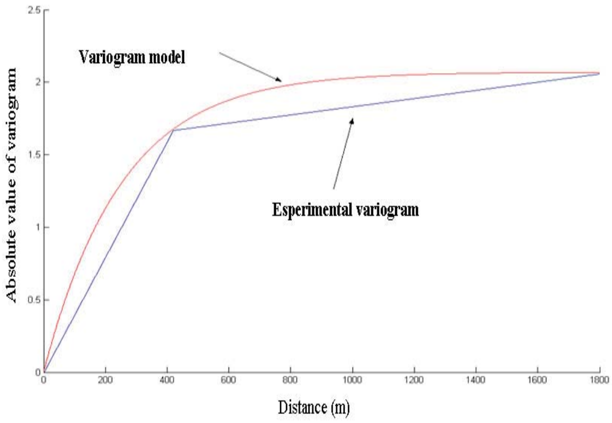

Figure 5. Representation of the exponential anisotropic model of Variogram.

The fig. 5 shows the trend of the experimental variogram and the model function nearer to it. It is important to notice that in this case, the arrangement of variogram around to sill (reached asymptotically), is netter than what seen in fig. 4.

4 RESULTS AND FINAL COMMENTS

The application of isotropic and anisotropic models are performed for retrieving data of the City of Maglie [9] [10]. Table I illustrates data measured by deposimeters located in different areas included Maglie for terrain contamination with solvents [11] while Table II depicts data obtained after having applied the algorithm in the case of anisotropic modelling and algorithm [12] [13]. Making a comparison with Table I, we can see that the missing data for Maglie are

Sum of total PCB in ng/m² per day	Max deposition for total PCB in ng/m² per day from july 15,2009 up to june 30,2010												
	july-09	**aug-09**	**sept-09**	**oct-09**	**nov-09**	**dec-09**	**jan-10**	**feb-10**	**mar-10**	**apr-10**	**may-10**	**june-10**	**average**
Maglie	2,890	1,500	4,760	17,070	4,850	16,120	1,120	2,420	3,190	2,670	4,900	3,790	**5,440**
Cutrofiano	2,030	0,210	1,780	0,900	1,670	51,760	1,870	1,120	1,250	0,550	0,990	2,170	**5,525**
Gallipoli	3,100	2,640	3,320	1,930	1,070	1,570	1,440	0,880	1,560	0,410	1,090	1,220	**1,686**
Guagnano	14,910	0,570	9,790	1,500	1,500	6,580	1,970	1,030	5,090	5,760	7,480	2,320	**4,875**
Otranto	2,400	1,770	1,300	1,400	3,090	1,390	0,790	0,790	0,480	0,300	3,040	3,520	**1,689**
Melpignano	0,890	0,890	3,910	3,690	26,540	12,220	0,980	1,310	5,780	1,450	2,180	0,860	**5,058**

Table I. PCB Measurements from deposimeters

Sum of total PCB in ng/m² per day	Estimates of max deposition for total PCB in ng/m² per day from july 15,2009 up to june 30,2010												
	july-09	**aug-09**	**sept-09**	**oct-09**	**nov-09**	**dec-09**	**jan-10**	**feb-10**	**mar-10**	**apr-10**	**may-10**	**jun-10**	**average**
Maglie	*2,740*	*1,410*	*4,610*	*16,110*	*3,930*	*18,010*	*1,350*	*2,140*	*3,100*	*1,420*	*5,620*	*3,130*	***5,297***
Cutrofiano	2,030	0,210	1,780	0,900	1,670	51,760	1,870	1,120	1,250	0,550	0,990	2,170	**5,525**
Gallipoli	3,100	2,640	3,320	1,930	1,070	1,570	1,440	0,880	1,560	0,410	1,090	1,220	**1,686**
Guagnano	14,910	0,570	9,790	1,500	1,500	6,580	1,970	1,030	5,090	5,760	7,480	2,320	**4,875**
Otranto	2,400	1,770	1,300	1,400	3,090	1,390	0,790	0,790	0,480	0,300	3,040	3,520	**1,689**
Melpignano	0,890	0,890	3,910	3,690	26,540	12,220	0,980	1,310	5,780	1,450	2,180	0,860	**5,058**

Table II. Estimates for PCB in missing location (Maglie)

clearly overcome since Maglie, as location, is included in a grid (fig.4). The algorithm has shown that data from one location are obtained from the others. The degree of accuracy depends upon the variogram model. As consequence, this kind of algorithm will help the use of robot [14] for moving terrain containers for reclamation and predicting pollutant concentration on ground [15].

5. ACKNOWLEDGMENT

The authors are in debt with Vincenzo Cagnazzo for deposimeter data.

6. REFERENCES

[1] T. S. Virdee & N. T. Kottegoda, "A Brief Review of Kriging and its Application to Optimal Interpolation and Observation well Selection", Hydrological Sciences -Journal des Sciences Hydrologiques, 29,4,12/1984.

[2] Hein Putter and G. Alastair Young," On the Effect of Covariance Function Estimation on the Accuracy of Kriging Predictors", Bernoulli, Volume 7, Number 3, 421-438, 2001.

[3] A. Lay-Ekuakille, A. D'orazio, M., De Sario, V. Petruzzelli, "Analysis of Spatial Data in Inversion Techniques for Microwave Remote Sensing: Use of Intrinsic Random Function of Order k", XXVIth General Assembly of International Union of Radio Science, University of Toronto – Toronto, Ontario, Canada, August 13 – 21,1999

[4] A. Lay-Ekuakille, G. Vendramin, A. Fedele L. Vasanelli, D. Laforgia, A. Trotta, "Kriging Code for the Treatment of Environmental and Electrical Parameters for Photovoltaic Panels", 6th IEEE - International Multi-Conference on SSD (SSD'09). Djerba, Tunsia. 23-26 march 2009.

[5] A. D'orazio, A. Lay-Ekuakille, F. Prudenzano, "Mapping of Remote Sensing Measurements with Intrinsic Random Function Order K", XVI IMEKO World Congress, Vienna, Austria, September 25-28, 2000

[6] L. Piper, A. Lay-Ekuakille, A. Vergallo, V. Pelillo, "A Novel Pseudo-Stationary Modeling of Pollutant Measurement Prediction from Industrial Emissions", XX IMEKO World Congress Metrology for Green Growth September 9--14, 2012, Busan, Republic of Korea

[7] G. Andria, A. Lay-Ekuakille, M. Notarnicola, "Mathematical Models for Atmospheric and Industrial Pollutant Prediction", XVI IMEKO World Congress , Vienna, Austria, September 25-28, 2000

[8] A. Lay-Ekuakille, A. Cataldo, A. Trotta, "ATOVS Temperature Measurement Retrieval Using Kriging Technique", Proc. SPIE Remote Sensing, Crete, Greece, September 2002.

[9] A. Lay-Ekuakille, V. Pelillo, C. Dellisanti, F. Tralli, " Sar Aided Method for Rural Soil Evaluation", SPIE2002 Remote Sensing, 22-27 september in Crete (Greece)

[10] A. Lay-Ekuakille, P. Vergallo, N.I. Giannoccaro, A. Massaro, D. Caratelli, "Prediction and Validation of Outcomes from Air Monitoring Sensors and Networks of Sensors", V International Conference on Sensing Technology , ICST2011, Nov 28th – Dec 1st 2011, Palmerston North, New Zealand

[11] A. Lay-Ekuakille, I. Palamara, D. Caratelli, F.C. Morabito, "Experimental IR Measurements for Hydrocarbon Pollutant Determination in Subterranean Waters", Review of Scientific Instruments, Vol. 84, pp. 015103-1 (8 pages) 2013,

[12]J. Murphy, K. Sabharwol, Ran, "The Indicator Variogram and the Simple Kriging Estimator: Useful Tools to Complement Lithologic Correlation in a Complex Fluvial Depositional Environment", AAPG Bulletin, Volume 79, 1995.

[13] A. Lay-Ekuakille, P. Vergallo, F. Serrano, R. Velasquez, A. Trotta,"Predicting PV Module Characteristics with Outdoor Measurements: Modeling Improvements", IEEE I2MTC 2011, May 10-12, Hangzhou, China

[14]A. Massaro, F. Spano, A. Lay-Ekuakille, P. Cazzato, R. Cingolani, A. Athanassiou, "Design and Characterization of Nanocomposite Pressure Sensor Implemented in Tactile Robotic System", IEEE Transactions on Instrumentation & Measurement, Issue 8, vol. 60, pp 2967 – 2975, 2010.

[15] A. Lay-Ekuakille, A. Trotta, "Predicting VOC Concentration Measurements: Cognitive Approach for Sensor Networks", IEEE Sensors Journal, vol.11, issue 11, pp.3923-3030, 2011