Sequential Gaussian Simulation as a tool to improve PM_{10} sampling scheme in industrial sites.

Guido Alfaro Degan[1*], *Gianluca* Coltrinari[1] and *Dario* Lippiello[1*]

[1]Engineering Department, University of Roma Tre, Via della Vasca Navale, 79, 00146 Rome, Italy.

Abstract. This paper is focused on the role of spatial and variographic analysis in the phase of sampling design. In particular, when dealing with environmental variables such as airborne dust concentration all over a selected domain, the best approach to catch the spatial structure of the variable itself, implies the full and more detailed coverage of the domain. In this study this goal is achieved by means of about fifty airborne dust concentration field surveys all over a square area 350 mt wide in a quarry plant in the center of Italy. These data, sampled according with a regular pattern, did not allow to catch the spatial structure of the variable itself thus avoiding a satisfactory variographic modelling. To improve the sampling scheme an infilling procedure was performed by adding an increased number of samples. The selection of these further samples, less than 10% of the total amount, was carried out using sequential Gaussian simulations in those zones of the domain in which the highest variability was recorded. The final outcome shown a good result determining a good upgrade in terms of variographic modelling and final estimation at the cost of few further samples.

1. Introduction

Sampling strategy is a very important aspect of spatial analysis. The information deriving from raw data are in fact, deeply influenced by sample locations and their distribution in spatial domain. In particular, with fixed economic availability and with the consequent fixed number of potential samples, the quality of representation of variability structure of the variable is strictly dependent on the way the available information sources are located in the spatial domain.

That is the reason for which sampling strategy is so important in spatial analysis and its optimization is a key concept in such kind of applications. This paper aims to show an application of variographic instruments as a tool to check the influence of the sampling pattern on the variability structure and thus propose the best strategy.

Many sampling airborne dust approaches in quarry environment have been defined in literature, ([1, 2, 3, 4, 5]), and each of them is reasonably applicable in different conditions. Nevertheless, a generalization of an optimal choice is proposed focusing on the capacity of applying it conveniently in any situation.

As known, each sampling strategy is opportunely applied in any well defined situation, but all the methods are grouped into two main divisions that are systematic strategies and random ones.

As easy realizing, the first group concern all the methods based on regular division of the spatial context and the systematic allocation of samples, while the second one regards a random distribution of information sources.

In this study the attention is focused on the first group characterized by a systematic allocation of airborne dust samples.

The main feature of such kind of methods is the capacity to cover uniformly the whole domain of investigation, and to explore with details the elements of the system that are only partially known.

Among systematic ones, the transect sampling strategy is focused on the idea to investigate some variable that is expected to vary principally along a certain direction, while the isovalues sampling method is usually used for the vectorialization of raster map, for which the aim is to create a punctual set of isovalues for a certain variable.

The common feature of such methodologies is the lack of relation between the variability of the variable and the choice on samples allocation. In pure systematic sampling, only the resolution of the grid is defined before (usually as function of economic availability), with no consideration on eventual over- or under-estimation of the field, while the other two methods make some assumption on the behaviour of the variable, but with no precise knowledge.

The main problem regarding the bias introduced by sampling strategy regards the relationship between the scale of variability of the field and the scale of the exploration grid. When the frequency of the sampling method is so precise, the information at that scale are redundant, while the ones at other scales are scarce or totally missing. In this last case the spatial analysis of sampled data reveals some lack in the coverage of the whole spatial structure of the variable. Such lack is often

* Corresponding author: dario.lippiello@uniroma3.it

related to the sampling strategy that is not able to catch entirely the behaviour of the field.
In a more direct way, observing a lack of information in the variographic analysis, the most important goal is oriented to infill sampling pattern with few additional samples. The best selection of these further locations represents the focus of this study.
Many authors, [6-7], dealing with other environmental variables spatially defined, suggested the selection of further surveying points taking into account those locations in which the standard deviation and the mean variance of the Kriged values is observed.
This application shows results of this idea applied to airborne dust concentration.

2. Materials and methods

Field tests were carried out in a basalt quarry in the southern part of Rome.
The plant, operating a single 8-hour shift per day, 5 days a week, consists of two different areas: the extracting area and the post processing plant.
In the first one raw material is shot by means of explosives while in the second one, the extracted rock is processed and divided according with different crushers and sieves.
The monitoring programme was developed in the processing area where dust sources revealed more intense and the airborne concentration variable more consistent, [8]. The devices utilised measured both the environmental PM_{10} concentration and the total suspended particulate (TSP). Samples consisted both in direct readings and gravimetric ones for a total amount of 44 values divided in two groups: 36 sampled according with a rectangular pattern and 8 surveys aimed to define local variations and selected in a second phase. PM_{10} concentrations were measured using a TCR Tecora Bravo Plus M sampler with a specialized sampling head which conforms to the UNI EN 12341 specifications. The particulate was collected on cellulose nitrate filters of 47 mm in diameter. Gravimetric measurements were then carried out on the filters both before and after sampling. This analysis was conducted by means of an analytical balance (Exacta series ABT 120-5 DM) with a sensitivity of five decimal places. Before carrying out the weighing operations the filters were prepared in a drier so as to reduce any error due to the hygroscopicity of the instrumentation or filters themselves. Each value was determined on the basis of three independent but consecutive weightings. With regard to the reliability of the measurements, the effects of experimental uncertainty were considered to be linked to two factors: firstly in connection with determining the deposited mass and secondly regarding the uncertainty due to the volume of air captured in the sample. The first aspect was quantified through blank filter analysis, whereas the second was quantified following the calibration of the air inlet sampling pumps by means of a bubble primary flow calibrator accredited by the United Kingdom Accreditation Service (UKAS) and calibrated both for high and low flow rates and the associated random error (within ±1% error). The combination of the two effects cited above allows the total uncertainty associated with the measurement to be estimated within two standard deviations.
Direct readings were carried out by means of photometer that is a real time detector based on the optical properties of the inlet particulates. Once airborne dust is collected into the sampling chamber, the device returns its concentration in a proportional way to the scattered light beam. The device has, in fact, an internal sensor detecting scattered light that also generates a current pulse proportional to many factors the most important of being particle size and distribution, density and refractive index ([9], [10]).
The following fig. 1 shows an aerial view of the selected area in which white/red points represent the sampled grid.

Fig. 1. An aerial view of the plant area and sampling points.

As shown, a rectangular pattern was drawn to characterize a square zone 350 mt. wide.
The 36 samples (white points) covered in a homogeneous way the plant processing area and constituted the input data set.
According to a geostatistical approach, each single value of the regionalized variable (in this case airborne dust concentration), and above all its spatial variability structure, is assumed to be described efficiently in a stochastic framework. Therefore, from a probabilistic point of view, any spatial variable z(x), measured at a specific location, is the outcome of a random process that generated it from a random variable Z(x).
A first and easy kind of approach to better understand the application of geostatistical methodology is the empirical one. Let us consider a certain spatial variable z(x) measured in n some locations x_i and some distance vector h, oriented in a certain direction, with an associated tolerance Ɛ. Let us consider the groups of pairs of samples that have a distance k h (with k as 1, 2, 3, etc.), along that direction, within Ɛ tolerance and calculate for each group the scatterplot and the correlation coefficient.
Now, if we substitute the standardized covariance (correlation coefficient), with the variance of histograms

of pairs for each class distance, we obtain a function that shows some peculiar differences.

Variance of increments, that can be regarded as variogram function (more precisely as twice the variogram function), presents a mirrored shape of covariance one. It is zero at zero distance and increases with increasing distance tending asymptotically to a certain value that is usually comparable to the sample variance of the whole dataset. This last assumption may be accepted when the first moment is spatially invariant and covariance itself does not depend on position but only on h distance between two selected points.

So variographic function is one of the most important tools of the structural analysis and just the transposition of such empirical approach. It is defined as follows:

$$\gamma(x,h) = \frac{Var\left[z(x+h) - z(h)\right]}{2} \tag{1}$$

Often, covariance function cannot be computed because of its lack, in the case the mean of the variable is not constant. Contrarily, variogram of increment of the variable, is an universal good estimator of the spatial variability that can be correctly used also when the variable is intrinsic. Besides, the shape of variogram curve helps us to catch the non stationarity degree of the field, in order to choose the more appropriate method to process it. Once the variogram has been computed, if its shape is revealing enough stationary to proceed with classical methods, the next step is the fit of the theoretical variogram function. In many cases the raw data and their related experimental variogram cannot be exploited to obtain a complete knowledge of the variability structure of the phenomenon, and some species of continuous function is needed. The inference of the experimental variogram is based on a least square fit of the experimental values of semivariance for each lag. Such fit can be made by means of the so called authorized models. The authorized functions are the ones that follow particular conditions, among which the main is to be positive definite.

By fitting a continuous mathematical function on raw variogram we can exploit such powerful instrument in order to model the variability structure for the whole spatial domain (not only on the points we have measured values). Such precious information can be used in estimation and simulation processes.

So starting from raw data, a variographic analysis was carried out under different conditions.

Result is shown in fig. 2

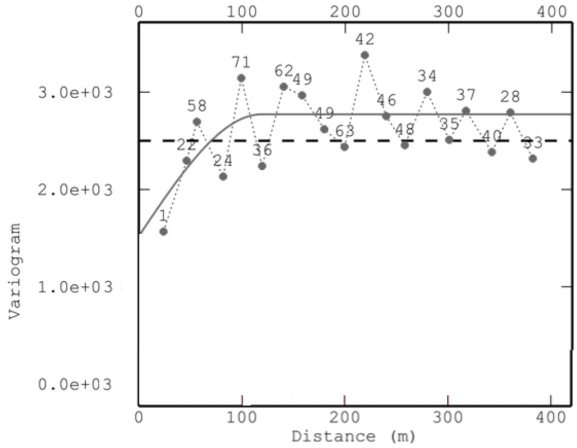

Fig. 2. Variogram of raw data.

in which black points represent the pairs at a selected distance (abscissa) while the continuous line represents the analytical function best fitting raw data.

On analysing the variogram in fig. 2, the most important outcome is represented by the high variation observed at the small scale. This last, in general applications, and above all when dealing with environmental variables is assumed unreliable. In fact, except that for particular applications, environmental parameters are assumed to vary smoothly, thus, the non zero value of the nugget effect, rather than be constituted by analytic variance, can be explained only by the sampling strategy that has been not able to correctly modelling the uniformity of the variable.

The main goal is now represented by reducing analytical variance with Sequential Gaussian simulations.

The first step consist in mapping airborne dust concentration by means of a geostatistical estimation process. The Kriging method is applied to draw estimated airborne concentration maps.

Under stationary conditions, inference of the experimental variogram is performed and based on a least square fit of the experimental values of semivariance for each lag. Such a fit can be made by well-known continuous functions such as authorized models. Thus spatial variability is modelled and estimation may finally be carried out. The passage from a discrete information description to a continuous description is performed with the following linear estimation:

$$z(x_0) = \sum_{i=1}^{N} \lambda_i z(x_i) \tag{2}$$

In which λ_i coefficients represent the solutions of the so called Kriging system shown in the following equation (3):

$$\begin{cases} \sum_{i=1}^{N} \lambda_i \gamma_{\alpha,\beta} + \mu = \gamma_{\alpha,0} \\ \sum_{i=1}^{N} \lambda_i = 1 \end{cases} \tag{3}$$

where $\gamma_{\alpha\beta}$ represent the variogram function computed for the α,β pair and μ the Lagrange multiplier.
The approach is an iterative cycle and assume a computation of Sequential Gaussian simulations (SGS) by means of the logical flow chart shown in fig. 3, which summarizes the logical steps of the process.

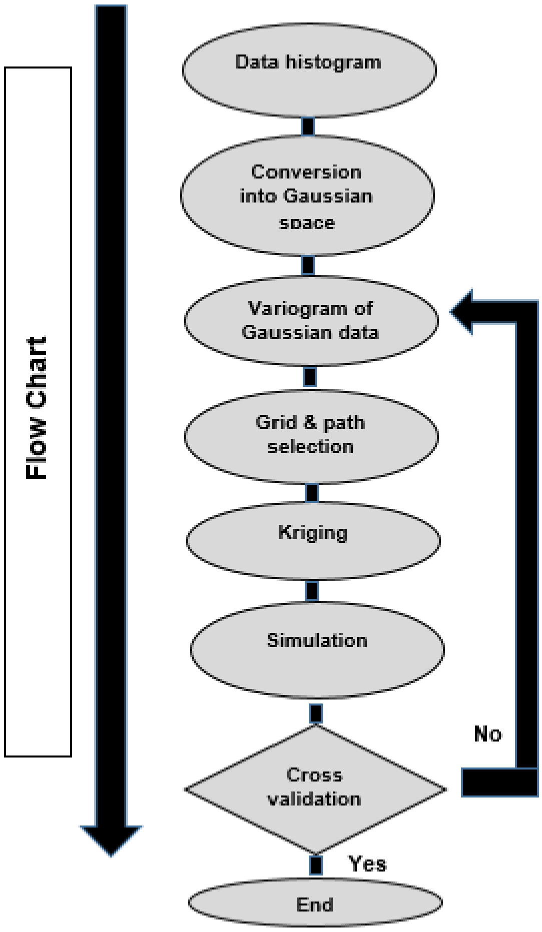

Fig. 3. Logical flow chart

The basic idea consists in the identification of additional samples located in those cells that show averagely the highest variability ([11], [12]).
The selection of such samples is performed from the kriged map and if cross validation shows good results they are addicted to the original dataset and eventually the computation of experimental variogram of the modified dataset is completed.
Each addition implies the assessment of the varying nugget variance which is compared with the analytic one. The process is assumed to be satisfactory when the two values converge [13].
In this case study, however, raw data are available and directly deriving from field tests so the new samples can be collected directly on the site, and not extracted by kriged map. In this way the sampling strategy is regarded as dynamic, composed by a preliminary exploration step and a further infilling with additional samples. Obviously the synopticity of the sampling process becomes a driving factor. In the case of the extractive site, 100 Sequential Gaussian Simulations have been calculated on PM_{10} variable, and the area (the grid cell) with maximum standard deviation of the 100 simulated maps are highlighted.
To examine the most significant results of this procedure, in the following figures 4, 5 and 6 are shown respectively the map of the selected area, the map of mean values resulting from the 100 simulated map and the map of mean variance.

Fig. 4. A view of the domain

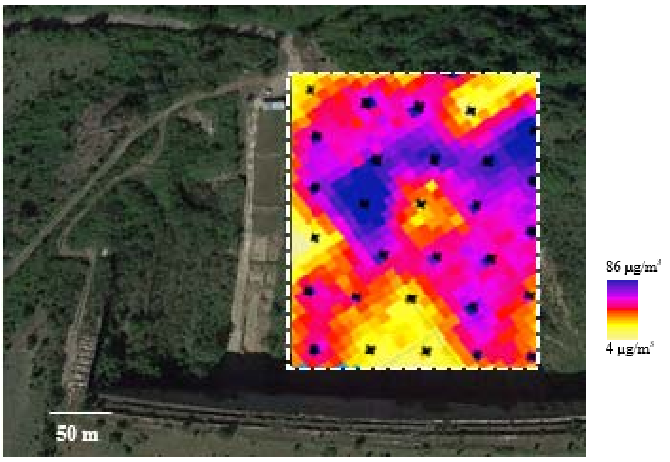

Fig. 5. PM_{10} mean concentration map

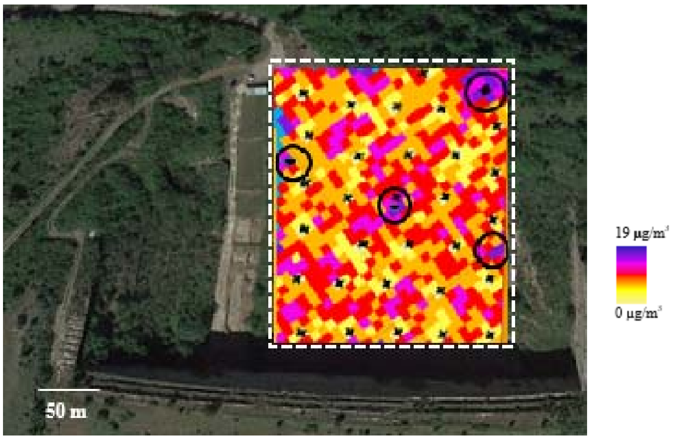

Fig. 6. PM_{10} mean estimation variance map

The four zones marked with black circles in figure 6 correspond to those areas in which the estimation variance returns highest values. The range goes to 15 to 19 µg/m^3.
In such locations further data were collected in the second stage of the field campaign for a total amount of eight readings corresponding to two samples for each mini area.
These further samples were added to the data set and the corresponding variogram was plotted.
The resulting output is shown in fig. 7.

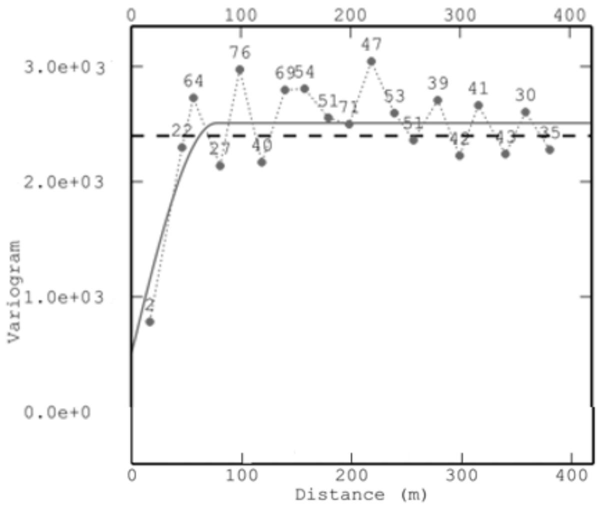

Fig. 7. Variogram of increased data set.

In the following table 1 the most important result are summarized. In particular the nugget variance is put in comparison with the corresponding infilling samples.
At a first sight the most direct outcome is the decreasing tendency of nugget variance when adding few samples. Its value, in fact, tends to reduce by a factor of two with only four additions.

Table 1. Nugget variance and corresponding samples.

# sample	Nugget variance
37	1.4 10^3
38	1.1 10^3
39	9.2 10^2
40	6* 10^2

3. CONCLUSIONS

The assessment of airborne PM_{10} concentration constitutes an important step towards examining the environmental compatibility of the quarrying plant with the surrounding area but also a basic tool to assess workers' exposure in the plant itself. In this regard the EU directive and the relative Italian legislation (Italian Decree 81/2008) require periodical checks to ensure that the established standards are upheld throughout the operation and development of the plant.
The procedure described for the infilling of an existing sampling plan wants to represent an indication on how to rebuild the sampling strategy for successive surveys.
From an analytical point of view, after the phase of variographic analysis and computation, the main focus is reducing the nugget variance of the modelled variogram. The aimed result is represented by the analytical variance whose value may be considered as about 10% of the median value of the raw data.
So, after the sampling stage which is generally carried out according with a systematic strategy, the selection of few further readings, collected in specific zones, has determined an increased quality in the estimated map.
These further samples were collected in those zone revealing the highest estimation variance after a fixed number of simulation. Moreover, these samples, in this field test not exceeding the 10% of the total amount of data, are compatible with economic availability and cost, brought an important improvement towards the correct coverage of spatial variability.
Finally this application proved to be a useful tool also taking into account that airborne samples generally imply long surveys as a satisfactory dust mass has to be collected on the filter at a fixed flow rate.

References

1. Vella A.J., Camilleri R. Fine dust emission from softstone quarrying in Malta. Xjenza 10, pp 47-54, 2005.
2. Jones, T., Blackmore, P., Leach, M.. Characterisation of airborne particles collected within and proximal to an opencast coalmine: South Wales, UK. Environmental monitoring and assessment, 75, pp. 293-312, 2002.
3. Ukpebor, E.E., Ukpebor, J.E., Oviasogie, P.O., Odiase, J.I. &. Egbeme, M.A, Field comparison of two Total Suspended Particulates (TSP) samplers to assess spatial variation. International Journal of Environmental Studies, 63 (5), pp 567-577, 2006.
4. Petavratzi, E., Kingman, S. W., Lowndes, I. S.. Assessment of the dustiness and the dust liberation mechanisms of limestone quarry operations. Chemical engineering and Processing, 46, pp. 1412-1423, 2007.
5. G. Alfaro Degan, D. Lippiello, M. Pinzari. Effectiveness of airborne dust control strategies in opencast quarrying activities: A case study near Rome. GEAM, **139**,2 pp. 5-12, (2013)
6. O. Asghari,, F. Soltni, H.B. Amnieh. The Comparison between Sequential Gaussian Simulation (SGS) of Choghart Ore Deposit and Geostatistical Estimation Through Ordinary

 https://doi.org/10.1051/e3sconf/201912809009

Kriging. Australian Journal of Basic and Applied Sciences, 3(1), pp. 330-341 (2009).

7. Vann, J.,Jackson, S., and Bertoli, O. 2003 Quantitative kriging neighbourhood analysis for the mining geologist - A description of the method with worked case examples. 5th International Mining Geology Conference. pp. 215-223.K.Y.K Chung., N.P. Vaughan. Comparative laboratory trials of two portable direct reading dust monitors. Annals Occupational Hygiene, 33, pp. 591-606, (1989).
8. G. Alfaro Degan., D. Lippiello, M. Pinzari. Field evaluation of PM_{10} detectors in a quarry environment. Int. Journal of Sust. Dev. Plan. **10**,3, pp. 361-372, (2015).
9. P. Kuusisto. Evaluation of the direct reading instruments for the measurements of aerosols. American Ind. Hyg. Assoc, 57, pp. 1013-1018 (2010).
10. C.J. Tsai, T.S. Shih, J.D. Lin. Laboratoty testing of three direct reading dust monitors. American Ind. Hyg. Assoc. 15, pp.48-56, (2000).
11. S. Khosrowshahi, W. J. Shaw, G.A. Yaetes. 2005. Quantification of the risk using simulation of the chain of mining - case study at Escondida copper, Chile., In Ore body modelling and strategic mine planning: uncertainbty and risk management models, Ed. R. Dimitrakopoulos, Spectrum Series Vol. 14, pp. 21–29, Carlton South, Vic., AusIMM.
12. J.F. Costa, J.A. Zingano and J.C. Koppe. Simulation - An Approach to Risk Analysis in Coal Mining, Exploration and Mining Geology, 9 (1), pp. 43-49, (2000).
13. G.G. Pilger, J.F.C.L. Costa, J.C. Koppe. 2005 Improving the Efficiency of the Sequential Simulation Algorithm using Latin Hypercube Sampling. In: Leuangthong O., Deutsch C.V. (eds) Geostatistics Banff 2004. Quantitative Geology and Geostatistics, vol 14. Springer, Dordrecht.