Studia Universitatis "Vasile Goldiş", Seria Ştiinţele Vieţii
Vol. 26 issue 4, 2016, pp.455- 461

APPLICATION OF GEOSTATISTICS FOR REMEDIATION PURPOSES IN POLLUTED SITES

Fatima-Zohra Benmostefa Largueche[1], Hamza Cheniti[2*]
[1] Department of Mathematics, University Badji Mokhtar of Annaba, B.P.12, Annaba, 23000, Algeria
[2] Département génie minier, Ecole Nationale Supérieure des Mines et Métallurgie Sidi Amar Chaiba (Ex CEFOS) BP N° 233 RP Annaba 23000, Algérie

ABSTRACT: The main objective of this study is to highlight the utility of geostatistics for polluted soil remediation purposes. The second objective of this project was more practical one. It consists on the identification of areas that should be subjected to remedial actions and also on deciding which contaminant needs to be considered when remediation processes are taken. To achieve the described objectives, a contaminated site has been studied and the following steps have been followed: The contamination concentration limits beyond which action needs to be taken to remediate the ground contamination, in which case it is important to determine the areas that should be subjected to the appropriate remediation measures. A presentation of a case study will follow. A brief site description is given. Next, a spatial analysis of the site has been carried out. It consists essentially of: Firstly a primary process of the data which means that histograms and an unprocessed representation of the pollutant's distribution has been plotted for each contaminant. Secondly a graphic presentation of the pollution using inverse distance weight (IDW) and Kriging interpolation technique is shown. This research is exploring the possibility of stochastic simulation to produce more realistic image of the data distribution, compared with kriging errors .Finally an assessment concerning Kriging is presented and a balance between the advantages and disadvantages of its uses is discussed.

KEYWORDS: Soil pollution; remediation; inverse distance weighting IDW; Kriging; geostatistical simulation.

INTRODUCTION:

A soil is said to be polluted when it contains an abnormal concentration of chemicals potentially hazardous to health, to the environment or on occasions to other targets.

The problem with polluted soils lies not only in the inherent risks to health and the environment, but also in the fact that this pollution condemns the future use of these lands.

Under the rapid pace of industrialization, and the increasing need for employment, countries was putting to exploit these lands without thinking about future use.

Nowadays, the urgent needs to protect the environment and the particular attention accorded to industrials waste prevention, lead tendencies toward green production and friendly environment growth, this push governments to adopt, develop and maintain appropriate policies concerning waste management which will certainly help to improve the present situation. An effective policy consists on firstly remedying the existing contaminated land and secondly preventing as far as possible the formation of any further contaminated land.

The magnitude of the problem is clear; the remained difficulties are now the restoration of these lands. It is obvious that before any piece of land can be used for a new purpose its old use has to identified. It is also important to identify the new use of the lands. Then the identification and the quantification of areas within contaminated land which require clean-up is required to enable the engineers to evaluate and reduce as much as possible the cost of any restoration or remediation(Benmostefa Largueche,2006).

For such purpose, it is essential to know the specific spatial distribution of the pollutants.

So, the objective here is to predict the spatial variation of one or more contaminants. In fact, the delineation of highly contaminated areas from areas of low contamination is necessary; since remedial action will only be applied to those areas where the average concentration of contaminants exceeds certain levels. A technique that would give a close approximation of the true spatial distribution of contaminants would therefore be very useful.

Sampling is usually conducted to determine the pollutant concentration at several points. however, we are not able to use these data directly to answer the land manager or regional planner who needs predictions of soil contaminants at unsampled sites, this is because of the multiplicity and complexity of the processes involved (Lark, 2012).

Many techniques are available to estimate the concentration of the contaminant at other points than the sampling points.

For instance, one of the easiest methods used to describe the distribution of a contaminant based upon a few data points is the linear interpolation. It consists on interpolating between data points using predetermined weighting function. Such weighting schemes may not reflect the actual distribution of the contaminant (Li J, 2008). Another technique commonly used is trend surface analysis. This method assumes that the data can be described by a polynomial surface where random error is assumed to be responsible for a deviation between the observed and estimated values. Such an assumption is not reasonable for interpreting contaminated sites, since the factors which are causing

** Corresponding author: CHENITI Hamza, Département de génie minier, Ecole Nationale Supérieure des Mines et de la Métallurgie, ex-CEFOS Chaiba B.P 233 RP Annaba, Algérie E-mail: hamza.cheniti@ensmm-annaba.dz*

a high concentration of a contaminant at a point are likely to cause the same effect at nearby points. A technique which takes into account the spatial correlation of the data is certainly more appropriate to describe the spatial distribution of the contaminants (Journel AG, 1978), (Oliver and Webster, 2014). A more detailed study of the problem shows that in certain ways the problems associated with describing the distribution of contaminants in contaminated lands are similar to the problems encountered in describing the distribution of mineral deposits.

RESEARCH METHODS:

The geostatistical estimation technique which may fulfil all the required criteria is called Kriging technique. This technique has been named after a mining engineer D.G. Krige who first applied it in 1951 in the mining field to estimate the average grade and total tonnage of ore reserve of the South African mines. Thus Kriging certainly overcomes some of the latest described problems. Kriging technique was primary developed to solve mining and geological problems but has since found through the latest few years wide applications in different fields such as groundwater, radiological, rainfall and medical applications. In a general way, this procedure can be implemented in all cases where some spatial correlation between sampling points is observed (Goovaerts, 1997), (Malvić and Đureković, 2003). As Kriging has already produced excellent results when used in the mining field, if it can be used to estimate the spatial distribution of contaminants in contaminated sites, the environmental engineers will possess a powerful tool for evaluating the exact sectors which require clean up. This will allow the reduction of the costs of remediation, which will help to make it more accessible than what it was in the past and thus more manageable and more affordable.

Introducing the site (Watford Waste Site):
The investigated site is located at Watford which is in the District (borough), County of Hertfordshire, England, situated on the north-west periphery of London. This site used to be a gas works site, and it is likely to have become contaminated with a wide range of chemical substances. The site is covered with hard standing and is to be redeveloped into a non food retail outlet. Wimpey Environmental Limited had provided a site location and tabulated results of the chemical analysis.

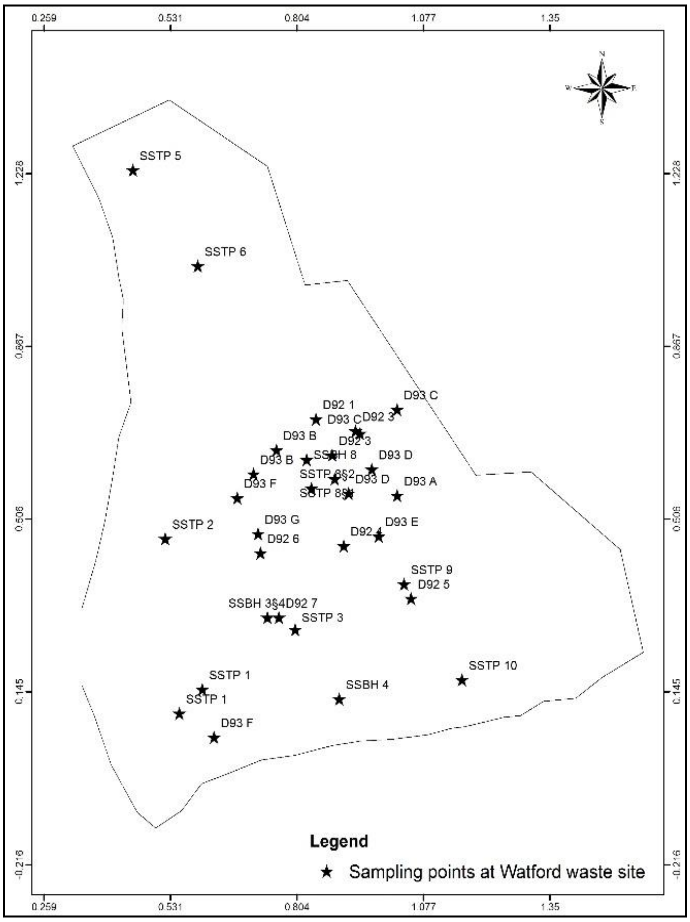

Fig 1 .Sampling points at Watford waste site

Spatial analysis: From the original tables, the contaminants can be divided in two categories. The first category will represent the contaminants whose concentration at sampling points is under the value of the trigger action limits.

The second category will include essentially the other contaminants that have at least one sampling point with concentration higher than the actual trigger action limits. In practice the contaminants contained in the first category should not be considered for the analysis of the sites, as we are interested in identifying which sectors of the site require cleanup. But from a theoretical point of view, as we are also interested in assessing the suitability of Kriging for analysing contaminated lands all the contaminants have been undertaken.

Results will be shown for pH, total chromium, total lead and elemental sulphur the latter three being significant contaminants. PH is not itself a contaminant as it is a measured parameter, but when presents in an certain level with other contaminants it can also be considered as a contaminant because of its effects. The spatial analysis starts by plotting a histogram for each variable to allow evaluation of its distribution, followed by a presentation of 2D Dots maps showing the real location of every sampling point for each contaminant and the real values of every contaminant at their specific points.

To carry out the spatial analysis, the 2D Dots maps are followed by a graphic presentation of the pollutions using the Kriging interpolation technique and the inverse distance wight interpolation technique (Shahbeik Sh, Afzal P, 2013). To allow this graphic presentation to be done, the three guidelines were used in deciding which contaminant needs to be considered when determining the areas that require cleanup. These levels will represent the legend for each contaminant, for both representations.

To terminate the analysis it is essential to decide which contaminant needs to be considered when determining the areas that require clean-up.

Before any statistical analysis of the data is done, it is important to first perform some elementary statistics. These elementary statistics are the primary process of the data. It consists of drawing the histogram for every contaminant concentration which provides the first information about the distribution of the probability law of every variable and secondly producing a 2D Dots map, for every contaminant which gives an overview about its real dispersion.

Before developing and modelling any experimental semi-variogram, a histogram or scatter plots (correlation diagrams) could be used, to check for outliers and non homogeneity. In fact the histogram is a valuable tool in determining whether the sample distribution is reasonably symmetrical and to detect visually possible outliers, or sample values which are abnormally high or low. However the shape of the histogram is affected by the limits of the classes used to group the samples. A histogram has been drawn for every contaminant. These histograms plots did not show a normal sampling distribution for any of the contaminants under study. From these histograms it was apparent that the data for every contaminant are not normally distributed. A lognormal transformation was then performed to try to normalise the data. None of the histograms of the transformed data has shown a normal distribution. So the non transformed data has been used to perform the sites spatial analysis.

The histograms showing the sample distribution of elemental Sulphur, Carbon Disulphide and Sulphate reflect a constant sample distribution which means that normally these three contaminants would not be considered for any further analysis. Despite the shape of their histogram it was interesting to analyze these three contaminants. The unprocessed representation of the pollutant's distribution consists on the 2D Dots maps, which have also been plotted for each contaminant. These maps give the actual value for every contaminant at the exact location of the sampling points and can only be plotted for an irregular dataset. A legend is automatically created and displayed. It can be modified at any time. Titles and axes have been added to make the maps more accurate and easier to follow. These maps were very useful. They have been used to appreciate the level of every contaminant at every sampling point.

Before starting using the Kriging interpolation process, it is assumed that the intrinsic hypothesis for every
contaminant is valid. So, it is assumed that the expected mean value of the contaminants in the regions of interest is constant. It is also assumed that there is no significant drift present in the data so that the point Kriging program contained in ESRI ArcGis could be used (ESRI, 2017). Isotropy was also assumed for the sites and has been checked by developing semivariograms in the four main directions using the ESRI ArcGis. In order to give a graphic presentation of the distribution of the pollutants using Kriging interpolation process, a semi-variogram has been plotted for every contaminant as it is the first stage of Kriging. After selecting the relevant parameters, a theoretical model has been chosen between the models provided by ESRI ArcGis (ESRI, 2017). to fit these experimental semi-variograms (Oliver and Webster, 2014).

In selecting the angle parameters a wide angle has been used to capture as many points as possible in an attempt to improve the fit of the semi-variograms and to avoid the lack of correlation where a default angle was used.

Despite the intrinsic hypothesis, the linear model has been implemented several times as a reasonable approximation since it has been seen that no other transition models could give a better fitting for the available data. Since h 0.7, near the origin both the spherical and the exponential models present a linear behaviour.

RESULTS AND DISCUSSION:

Our final objective is then to specify both the contaminants and their exact locations so that appropriate remediation plans can be undertaken.

pH: The value of pH is estimated to be higher than 08 both by the Kriging and IDW method. So since the level of alkalinity is estimated to be higher than the action trigger value, remedial actions need to be taken.

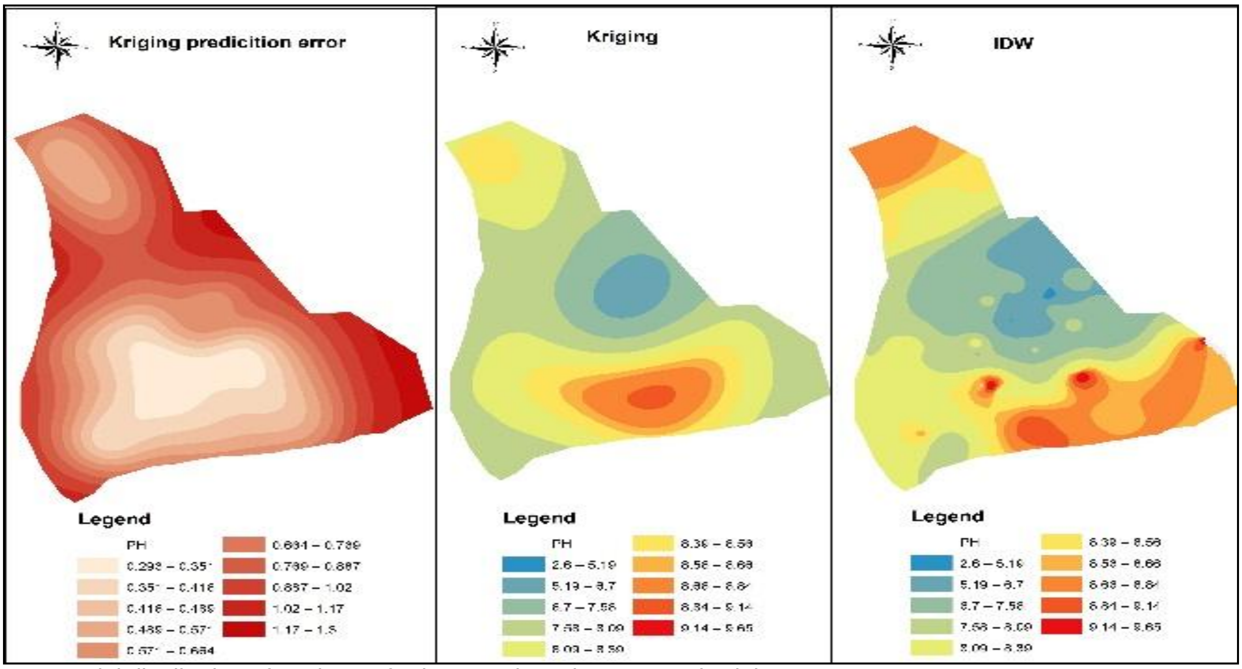

Fig 2 . Spatial distribution of PH in Watford waste site using IDW and Kriging.

Total chromium: In a particular region of the site (in the middle) the value of Chromium total is estimated both by the Kriging and IDW method to be higher than 800 mg/kg. As this level has been suggested as the action trigger level for this contaminant, total chromium needs to be considered for remediation purposes.

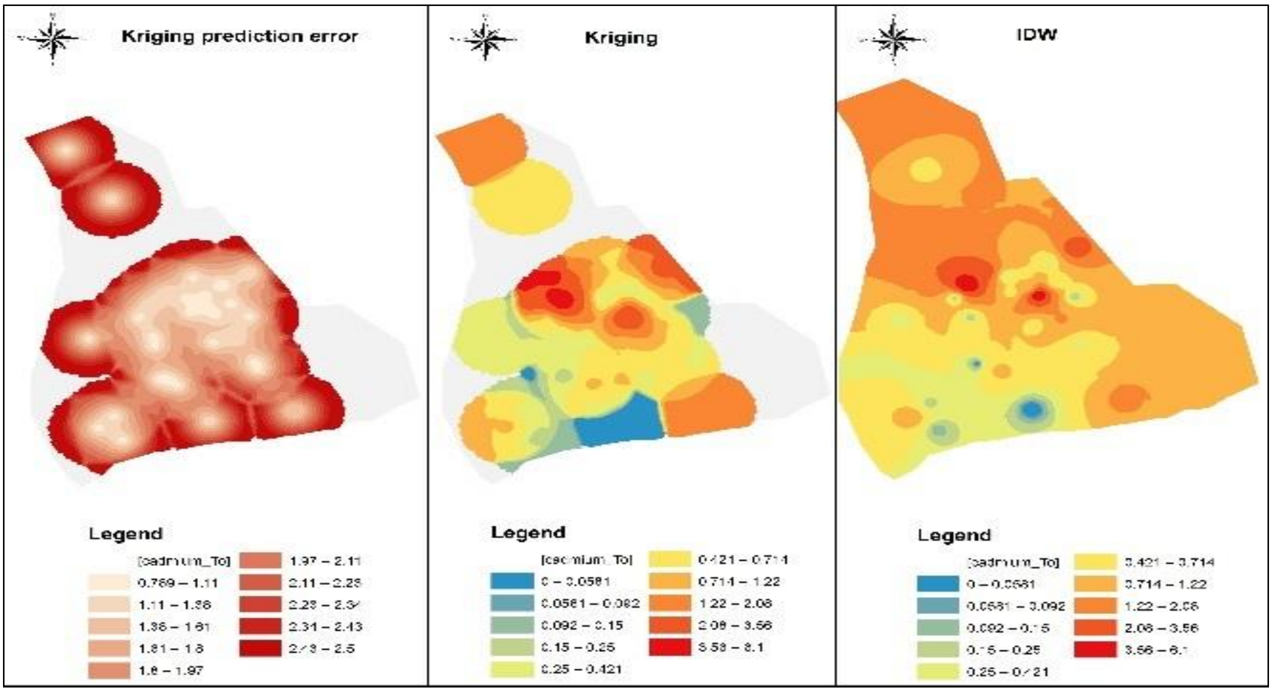

Fig 3 .spatial distribution of Total chromium in watford waste site using IDW and Kriging.

Total lead: The 2D contour maps obtained using Kriging technique shows that most of the site is typified by total Lead with a concentration higher than the 25 mg/kg the action trigger value. Since total Lead is considered as a source of danger, it needs also to be considered for a cleanup.

Comparing the figure 4 and 5 related to two realization of a gaussian geostatistical simulation of total lead, it appear that despite kriging is a good interpolation technique, it gives a image smoother than reality, On the other hand Geostatistical simulation (GS) generates multiple, equally probable representations of the spatial distribution of the attribute under study (figure .5). These representations provide a way to measure uncertainty for the unsampled locations taken all together in space, rather than one by one (as measured by the kriging variance)(Robertson et al., 2006).

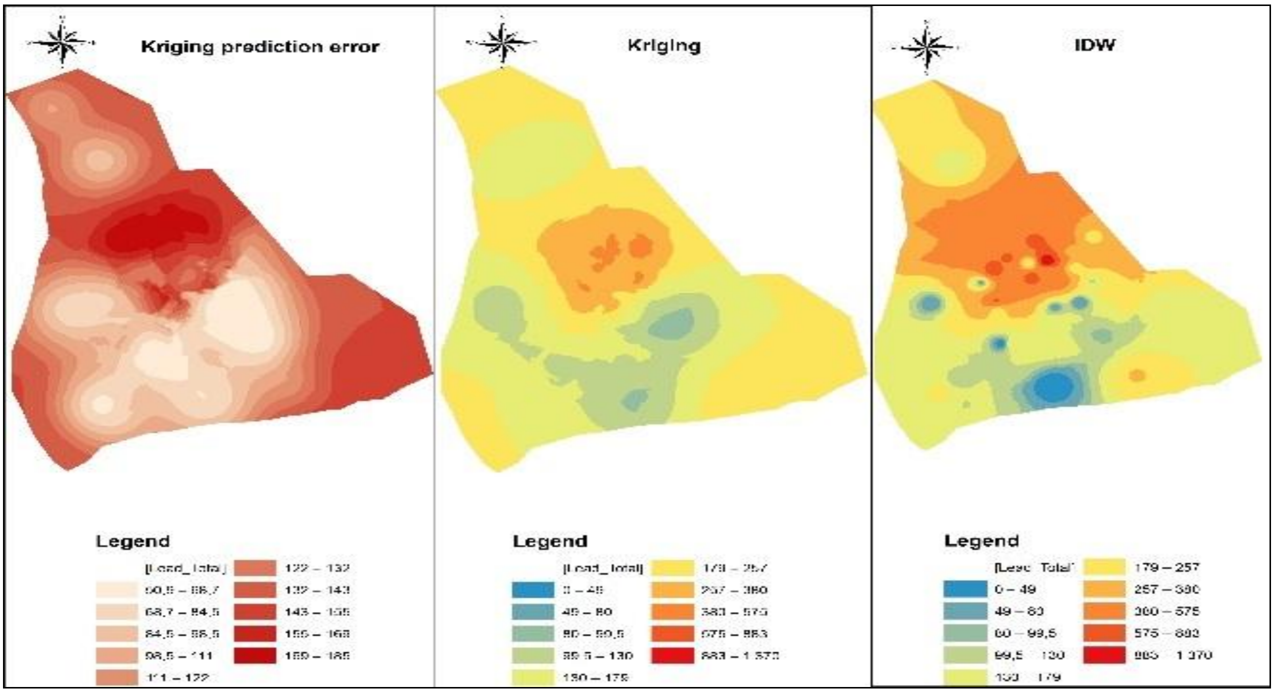

Fig 4 .spatial distribution of Total lead in watford waste site using IDW and Kriging.

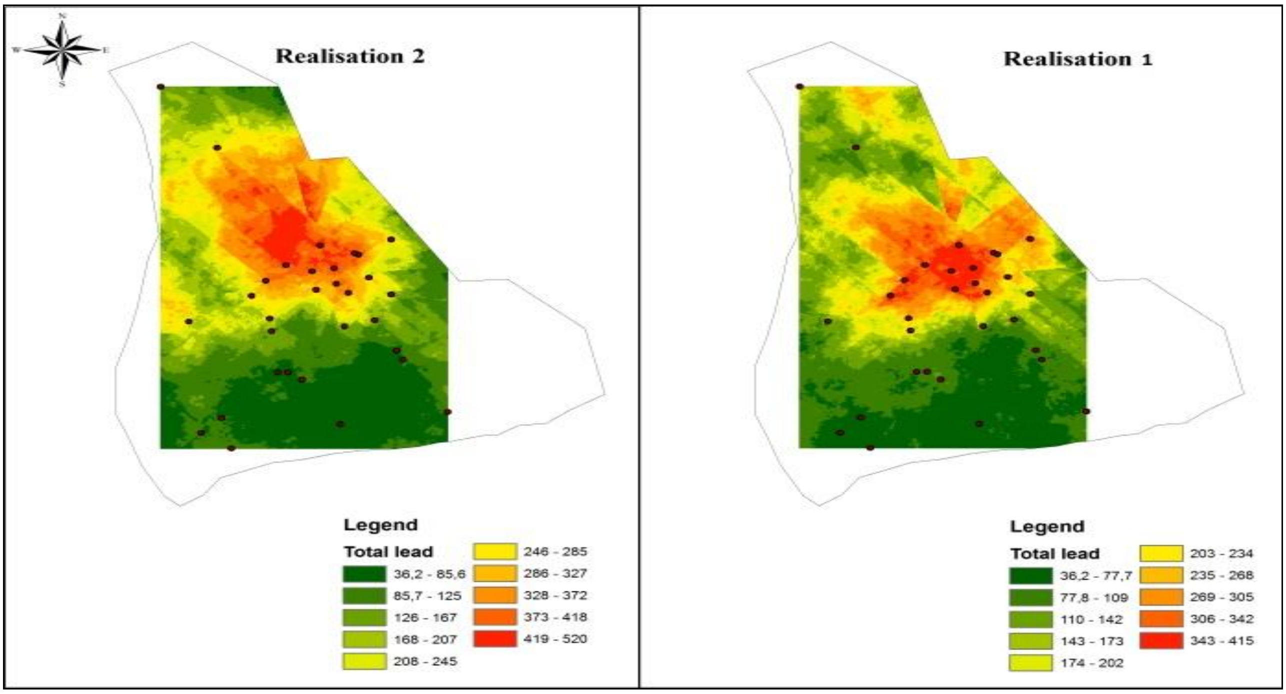

Fig 5 . Two realization of gaussian geostatistical simulation of Total lead in watford waste site.

CONCLUSION:

The results of the analysis of the three contaminants identified in the Watford waste site shows several comments regarding the applicability of Geostatistics for analysing contaminated lands.

The biggest advantage of the Kriging technique over many classical statistical procedures is that Kriging incorporates the spatial correlation of the data, while all the other classical statistical procedures do not. Another main advantage of Kriging over other interpolation techniques is its ability of quantifying the estimation variance, which will lead to define the precision of the resulting estimates. The standard error map can be used effectively in identifying the areas for which further sampling is necessary in contaminated sites. In fact, the error map shows the confidence envelope that surrounds the estimated surface. It expresses the relative reliability of the map of the estimates values. In areas of poor sampling, the error map will show large values, indicating that the estimates are subject to high variability. In areas of dense sampling the error map will show low values and at the sampling points themselves the estimation error will be zero.

Compared to the deterministic interpolation process, the Kriging method has further advantages.

While the deterministic interpolation technique estimates values of a contaminant in a sector where practically there is no sampling point so no data at all, the maps obtained by using Kriging estimation process show a red collar (see fig 4 kriging prediction error) which allows the technicians to know the zones where it is still necessary to sample. regarding the disadvantages of Kriging, Kriging techniques are based on a wide range of methods. These methods are derived from the regionalised variable theory. While compared with pre-existing techniques for analyzing the data they generally have got great advantages. In fact, they are based on strong theoretical basis. They also allow some estimation of the quality of estimates produced and they have some claim to statistical properties, as for example: unbiasedness, linearity and minimum variance.

On the other hand, these methods require certain strong assumptions to be made; assumptions which are rarely met in nature.

In fact, fundamental regionalised variable theory requires that at least the intrinsic hypothesis form of stationarity is true: local variations in the mean are accepted, but the semi-variogram must be necessary stationary over the entire area of interest. In practice, real data sets rarely even approach stationary. Universal Kriging and the generalised covariance method deal with non-stationary data, but even in these methods, the types and amounts of non-stationary are restricted to a few idealised situations. To solve this problem, most authors on Geostatistics suggest that assuming stationary is not so significant since local stationary is assumed; however there is no general proof of this and no statistical test to determine whether such an assumption is warranted: The fact remains that the theory of regionalised variable cannot be used under conditions where its defined form of stationary does not exist.

The pivot of parametric geostatistical methods is the semi-variogram which may be computed from the data set under investigation.

The experimental semivariogram obtained will often be different from all the theoretical models. Some skilful interpretations are required to fit one or more models to the empirical curve. It is also important to recognise breakdown of assumptions such as stationary, which will have direct distorting effects on the semi-variogram.

It is important to ensure that the semi-variogram computed from one data set will depend on this particular set. For other data sets the semi-variograms computed will be different.

The estimation of the semi-variogram may be difficult when there is a shortage of experimental points. This problem may be encountered in the case of a contaminated site, where ground conditions may prohibit access to certain areas of the site.

The choice of a technique has to be made. From the semi-variogram some departure from stationary has been diagnosed, it might be considered best to use universal Kriging or generalised covariance. On the other hand, if it is known that the data follow some complex non-normal distribution, it would probably be more appropriate to use disjunctive Kriging. When both situations occur, which should be then the recommended method? A method which will meet such a situation does not even exist and even if it combines the properties of both generalised covariance and disjunctive Kriging will be inevitably be too complex in terms of computing effort.

Generally, Kriging requires fewer samples than other spatial estimation techniques for obtaining an acceptable precision. Sometimes however, the number of samples required to estimate the semi-variogram may exceed the number of samples required to achieve a desired level of precision.

The Gaussian geostatistical simulation gives a more realistic image of the pollution distribution and allows uncertainty quantification for decision and risk analysis, a common question is then above which level of risk should we decide to clean a polluted area or develop specific land use policies.

The results presented herein show that further development on indicator simulation is needed. The concentration of a contaminant is likely to be highly variable; this technique may eliminate many of the problems associated with highly variant data sets.

Since integrating the estimated values with soil guide-lines values to map the risk from those contaminants is a vital step for determining which areas require cleanup, a more complete guide-line may be required.

The experience of dealing with contaminated sites will contribute in reducing the element of uncertainty at all stages of the reclamation of polluted land and a database compiling these experiences will allow practitioners in contaminated land to improve the quality of the remediation.

REFERENCES:

Benmostefa Largueche, F.-Z., 2006. Estimating Soil Contamination with Kriging Interpolation Method. Am. J. Sci. Appl. Sci. 3, 1894–1898.

ESRI, 2017. Environmental Systems Research Institute.

Goovaerts, P., 1997. Geostatisticsfor natural resources evaluation. Oxford University Press, New York;

Journel AG, H.C., 1978. Mining geostatistics. Academic press, London.

Lark, R.M., 2012. Towards soil geostatistics. Spat. Stat. 1, 92–99. doi:10.1016/j.spasta.2012.02.001

Li J, H.A., 2008. A review of spatial interpolation methods for environmental scientists. Geosci Aust 23.

Malvić, T., Đureković, M., 2003. Application of methods: Inverse distance weighting, ordinary kriging and collocated cokriging in porosity eval- uation, and comparison of results on the Beničan- ci and Stari Gradac fields in Croatia. Nafta 331–340.

Oliver, M.A., Webster, R., 2014. A tutorial guide to geostatistics : Computing and modelling variograms and kriging. Catena 113, 56–69. doi:10.1016/j.catena.2013.09.006

Robertson, R.K., Mueller, U.A., Bloom, M., 2006. Direct sequential simulations with histogram reproduction: A comparison of algorithms. Comput. Geosci. 382–395.

Shahbeik Sh, Afzal P, M.P., 2013. Ordinary kriging (OK) and inverse distance weighted (IDW) based on estimation error case study: in Dardevey iron ore deposit, NE Iran. (In press). Arab J Geosci.