The KTNVS Method

USING A VARIOGRAM SURFACE INSTEAD OF FITTING THEORETICAL MODEL VARIOGRAMS TO IMPROVE KRIGING

Erasmo Mejía

Universidad de Guanajuato, MEXICO
Norwegian University of Science and Technology
Trondheim, NORWAY
erasmo_mejia@stanfordalumni.org
emejia@quijote.ugto.mx

I.A.M.G. 2001

"… they would be more likely to realize that the parameters they are choosing for they random function-its expected value, its variogram and its stationarity (or lack of it)- are choices that they are making about their model and are not properties of the reality they are studying. Since the probabilistic model is figment of our imagination (a convenient and practical figment, but a figment nonetheless), its parameters are not real facts of the world."

Srivastava.

TABLE OF CONTENTS

TABLE OF FIGURES

The KTNVS Method. USING A VARIOGRAM SURFACE INSTEAD OF FITTING THEORETICAL MODEL VARIOGRAMS TO IMPROVE KRIGING

Erasmo Mejía

Abstract

This paper presents a bridge in applied geostatistics between supplying explicit experimental or implicit calculated variograms in block kriging. The proposed method: Kriging Technique with a New Variography Simulation (KTNVS) uses stochastic conditional simulation to calculate the spatial dependency expressed as a 3D autocorrelation surface. This surface maps the zones of high autocorrelation (zones of influence) and permits the user to avoid having to explicitly supply variogram models specifying isotropy/anisotropy as well as other complex or sophisticated features. A case study is presented where the estimation variance and estimation error are confronted using crossvalidation and the results are compared with those obtained by using traditional variogram model as input. This case demonstrates that the user is implicit not capable of capturing the complexity of the spatial autocorrelation if sufficient data are available, and a 3D autocorrelation surface approach is therefore preferable. The proposed method avoids the explicit variogram modeling used in traditional kriging approaches thereby avoiding the user skill dependant or model driven aspect of the estimation. The method therefore provides a user independent data driven approach to geostatistical kriging.

Introduction

It is still an every day usage in geostatistical reserves estimation to fit a model or theoretical variogram from experimental variograms as a prior step to kriging estimation. “The predominant practice has been first to calculate an experimental variogram and then to fit a model by trial and error with the help of computer software” (Englund and Sparks, 1988; Deutsch and Journel, 1992).

This fitting task has to relay mainly in two supporting points of application: 1. Human skill and 2. The sampling data. These points of application get critical when experimental variograms come out noisy or with a strong random component. Then variograms are difficult to interpret clearly to finding its basic characteristics. Therefore the fitting is not an easy task to do.

Human skill has to do with experience/intuition/ability, which is still better than others indirect. Automatic interpolation/smoothing and fitting methods have been applied and found non-satisfactory or not consistent. Maximum likelihood methods; the problem with these methods are that they rely heavily on the Gaussian distributional assumption and the estimates are biased (Cressie, 1991). The methods are unacceptable slow for sampling sizes larger than 150 (McBratney and Webster, 1986; Kitanidis, 1997). Ordinary least squares; the method however makes the unrealistic assumption that the differences are normally distributed, are independent of one another, and that all the estimated values have the same variance. The method has been rightly criticized to be unsuited for fitting of semivariogram models (Olea, 1999). Generalized least squares turns into an iterative procedure that does not always converge (Gotway, 1991).

Non-automatic fitting on the other hand, has to rely considerably on the knowledge/skill of the user. Different users will propose not necessarily the same model, turning this fitting stage into a users dependant one.

Sampling data involves different conditions to deal with. There are several sources of random components, they add up to make experimental variograms noisy, not so clear to interpret in their basic parameters nor its easy to see its own structures and shapes. Among all the different random components the one that affects more strongly is lack of data. Illustration of this: See appendix

TRADITIONAL APROACH

I. What do we expect from a variogram

The variogram, kernel of geostatistics, expresses the double meaning of Matheron's dual character random variables, and has its purpose in letting us see the structured as well as the random variance. Starting the random one after the area of influence which limits the first or structured one. To likely reaching the overall variance or sill thereafter. Expressed graphically as a flat curve, where direction is another of its qualities. The experimental variogram as a discreet realization of this random variable has the same objective too, and when it comes out noisy it is difficult to find out its area of influence, nugget effect, sill, etc.

After some understanding of the structures in the variogram, one has to figure out a known model that may possibly imitate what in the experimental variogram sometimes looks unclear, and some others may look complex. We explain them as hole effect, nested structures, isotropy, anisotropy, and so on.

This stage of interpreting the qualitative part of the variogram is no other than trying to find a match with the geology structures. To be able to find two dialects speaking of the same phenomena; structures spoken by the variogram and structures spoken by geology. We need to find this agreement between them in order to be certain to proceed to the quantitative part, the one where we do the calculation of the covariance matrix. When we fail in finding this agreement, but still fit parameters to a "blind" and unclear set of structures, then the quantitative part will be insecure and blind as well.

We need a mathematical model, an equation from where we can extract calculations to fill up the matrix, the covariance matrix, prior to kriging. When experimental variograms are complex in shapes, wiggling different forms, we want the theoretical variogram do the same, getting more complex too. We have models like the sine-hole effect, two or more expressions of it, cubic semivariograms, spherical, pentaspherical, several exponential models like gaussian, several power models, linear, prismato-magnetic, prismato-gravimetric, and several more.

II. Calculating the covariance matrix

Once a model is found, then the variogram is expressed as a vector function $\gamma(\boldsymbol{h})$, where the magnitude of $\boldsymbol{h}$ is the distance with a direction implied. Distance and direction where the random variable $Z(\boldsymbol{h})$ is analyzed in its spatial variation, expressed by:

$$\frac{1}{2}\,\mathrm{Var}[Z(\boldsymbol{x}) - Z(\boldsymbol{x}+\boldsymbol{h})]$$

This spatial variation describes explicitly the two aspects of Z(***h***), partially structural and partially random. A classical model of this is the spherical model shown below:

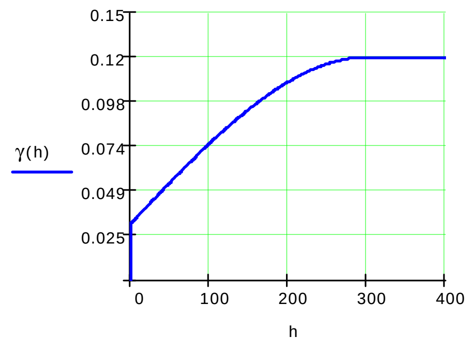

FIG-01 SPHERICAL MODEL OF A VARIOGRAM.

Here we are able to see the three basic parameters of the variogram, without forgetting to specify the direction it was calculated.

However, the discrete realization of the random variable some times does not come as clear as this model. Below we see Fig-02 where we find four experimental variograms calculated out of 90 drill holes with 1014 samples. The element is CaO.

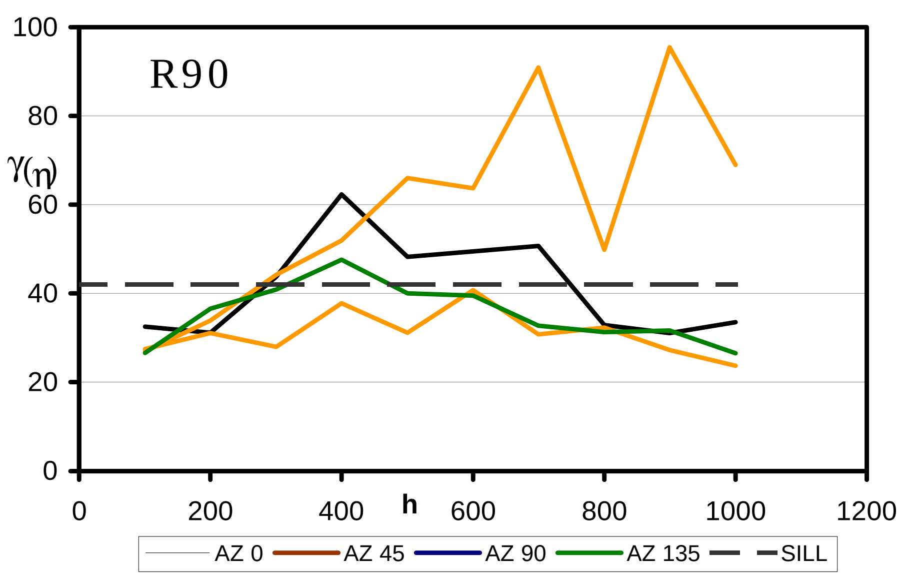

FIG-02 EXPERIMENTAL VARIOGRAMS IN 4 MAIN DIRECTIONS, DISTANCE IS IN METERS.

Trying to model these variograms we could find several different alternatives, and of course user's experience/skill will make the difference. There is no automatic nor deterministic method capable of define models from these set of experimental variograms. However there is a fit proposed, we can see it below in Fig-03

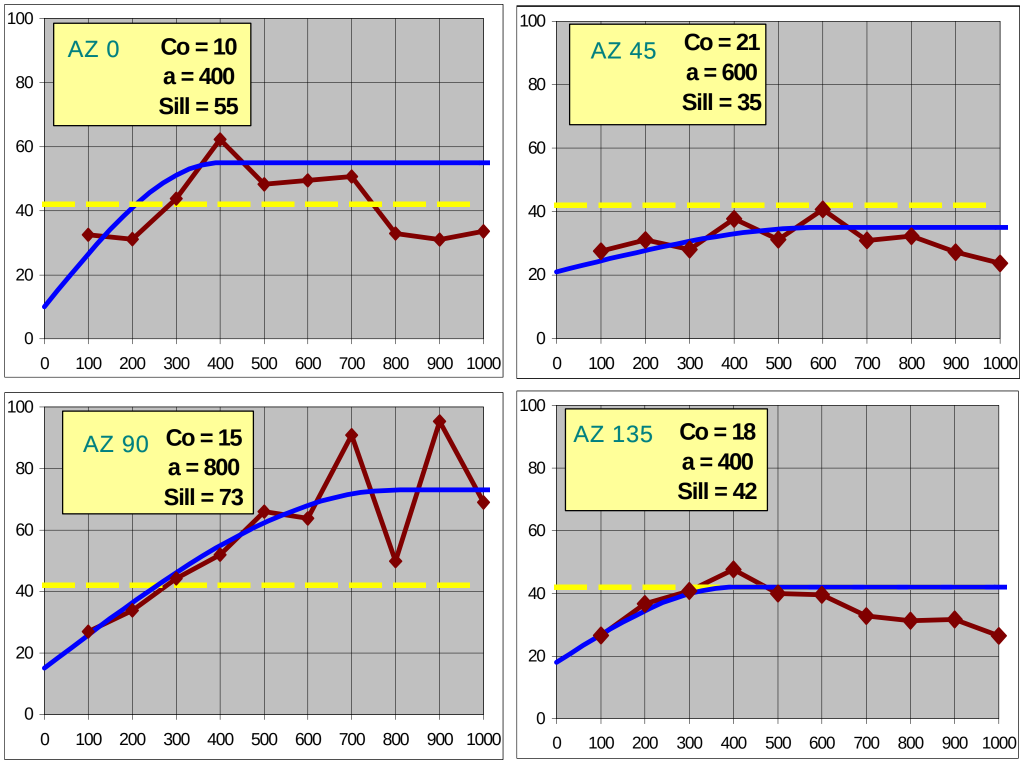

FIG-03 EXPERIMENTAL VARIOGRAMS AND FITTED MODELS. DISTANCE IN METERS.

Several questions are here to be asked, like: in AZ45 variogram would it be valid 600 m. for area of influence, or should it be interpreted as a pure nugget effect?. AZ 90 variogram should be fitted with a hole effect model. Two variograms point down below 20 in the nugget effect, while two others point over 20. The reader may find some more questions.

Continuing with the procedure, once we have the equation(s) of the model(s) to point or block kriging we select couples of samples within a neighborhood surrounding the block and depending on the direction of ***h***, from one sample to the other, in the couple, we select the corresponding variogram model and evaluate $\gamma(\boldsymbol{h})$. Repeat this until we finish either all the possible couples within the new neighborhood, or we reach a pre-established number.

III. Variogram Surface

What we are doing actually is mapping or evaluating a variogram surface defined by the model(s). When we are calculating $\gamma(\boldsymbol{h})$ in every possible direction as samples are placed in $\mathbf{R}^3$, then we are mapping a surface. The variogram surface of the 4 previous variogram models is shown below in Fig-04.

Even though not all of the programs or software available in geostatistics make use of 4 different models in its estimation process, let us however consider that possibility.

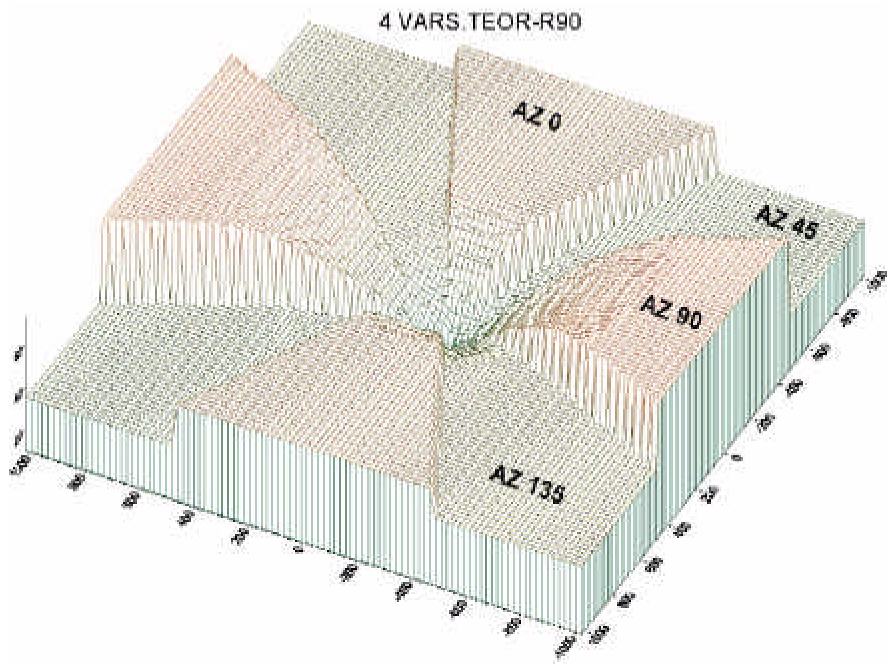

FIG-04 VARIOGRAM SURFACE FROM 4 THEORETICAL VARIOGRAMS.

Explicitly the procedure is then to evaluate or map this surface to obtain the entries of the covariance matrix for the kriging system.

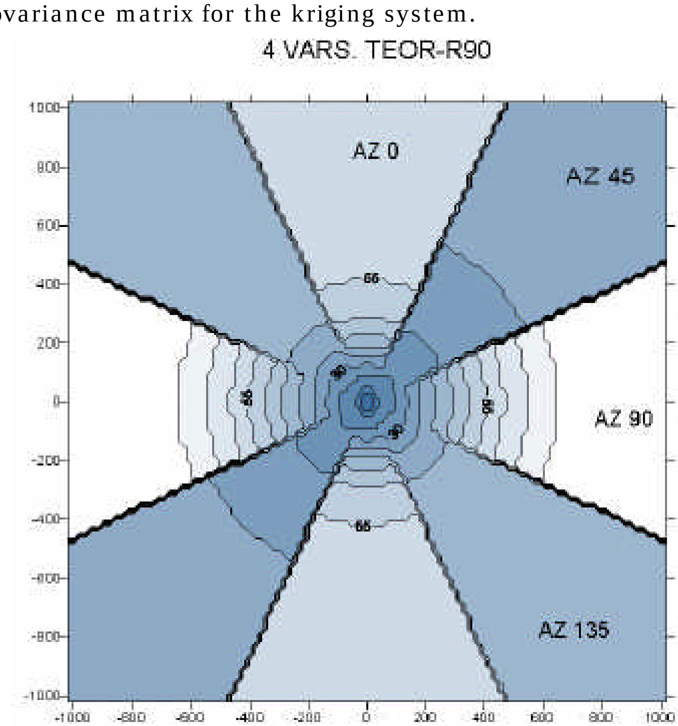

FIG-05 ISOVALUES OR CONTOUR MAP OF SURFACE IN FIGURE ABOVE.

The variogram surface from the experimental variograms is shown below.

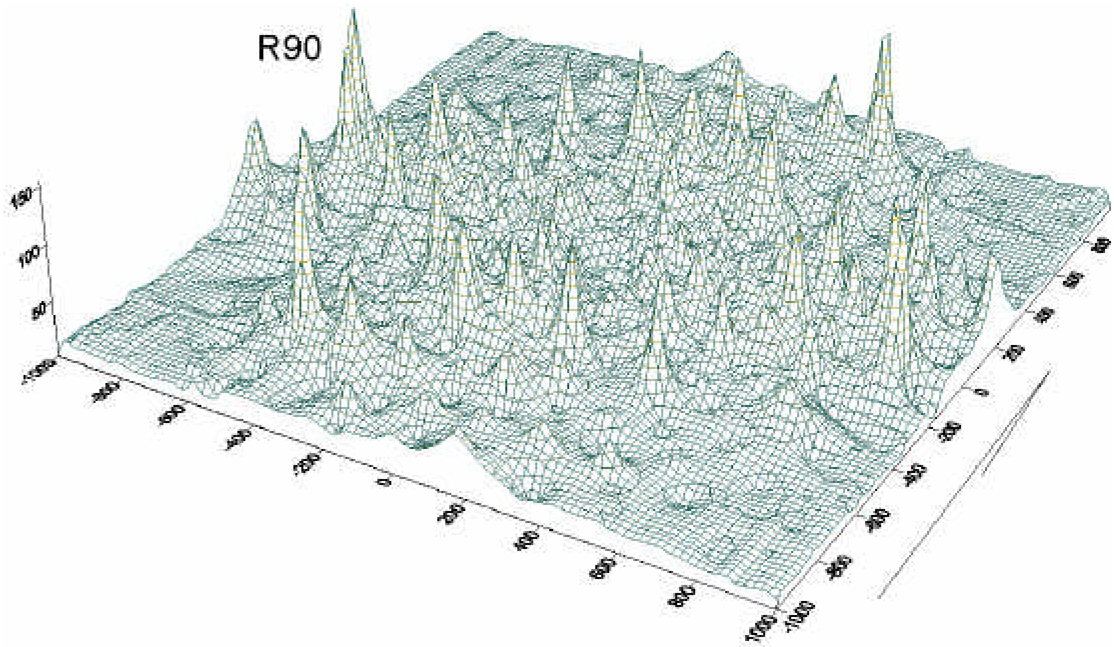

FIG-06 VARIOGRAM SURFACE FROM EXPERIMENTAL VARIOGRAMS R90.

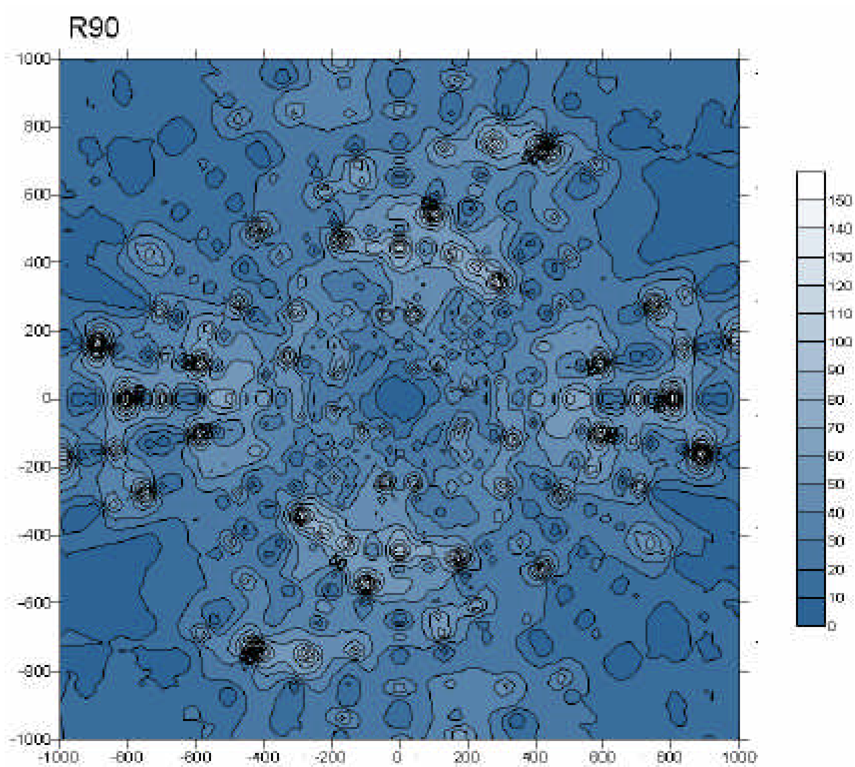

FIG-07 CONTOUR MAP FROM THE ABOVE SURFACE IN FIG-06

After these last two figures 06 and 07, we see clearly why we need to find a better representative of the experimental variograms. One reason is that variogram parameters are not clear enough from Fig-02 nor from Fig-06. Also because of the evaluation is necessary. We need to map values from a function. Then we have two good reasons why we look for models.

What would we do if experimental variograms would come out smoother and clean of noise, their structures clear and easy to interpret, should we still have to model them? Could we possible use themselves as empirical functions, map them and do the evaluation (interpolation) out of them to calculate $\gamma(\boldsymbol{h})$ and therefore the covariance matrix as we do with the models?

THE KTNVS APPROACH

IV. KTNVS method

The main problem here is lack of clarity in finding the characteristics of the variograms. That is why we make use of our skill/experience to figure out a model. Another possibility is drilling more holes to increase the sampling database trying to reduce noise. The case study presented here after an initial drilling of 60 holes (R60) with 814 samples, had a second campaign with 30 more holes and 200 samples, a total of 90 holes (R90) and 1014 samples. Please find a comparison between experimental variograms for both campaigns as well as other comparisons.

What most fitting techniques try to solve is the noise problem in terms of a representative, keeping the same sampling and therefore, the same experimental variograms too. KTNVS method starts from the beginning; It performs a simulation enriching the sampling to overcome the strongest source of noise in experimental variograms, when there is few sampling data (see appendix).

The figure below is the KTNVS variograms in its 4 main directions.

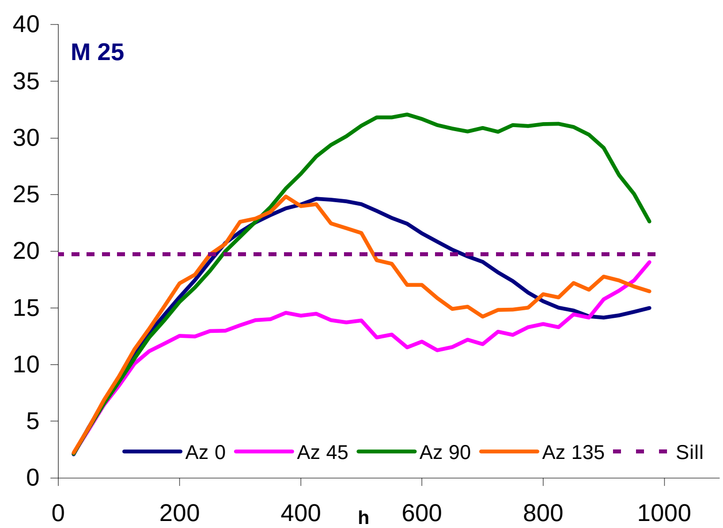

FIG-08 KTNVS VARIOGRAMS IN FOUR MAIN DIRECTIONS.

From the variograms above we can see several things: 1. Variograms are clean, not noisy 2. More clear structures 3. Consistency in Co or nugget effect. 4. No parameters to be estimated/guessed, and 5. No user dependence fitting.

Lets match now KTNVS variograms first with experimental variograms R90, then with geology. Figure below shows first matching.

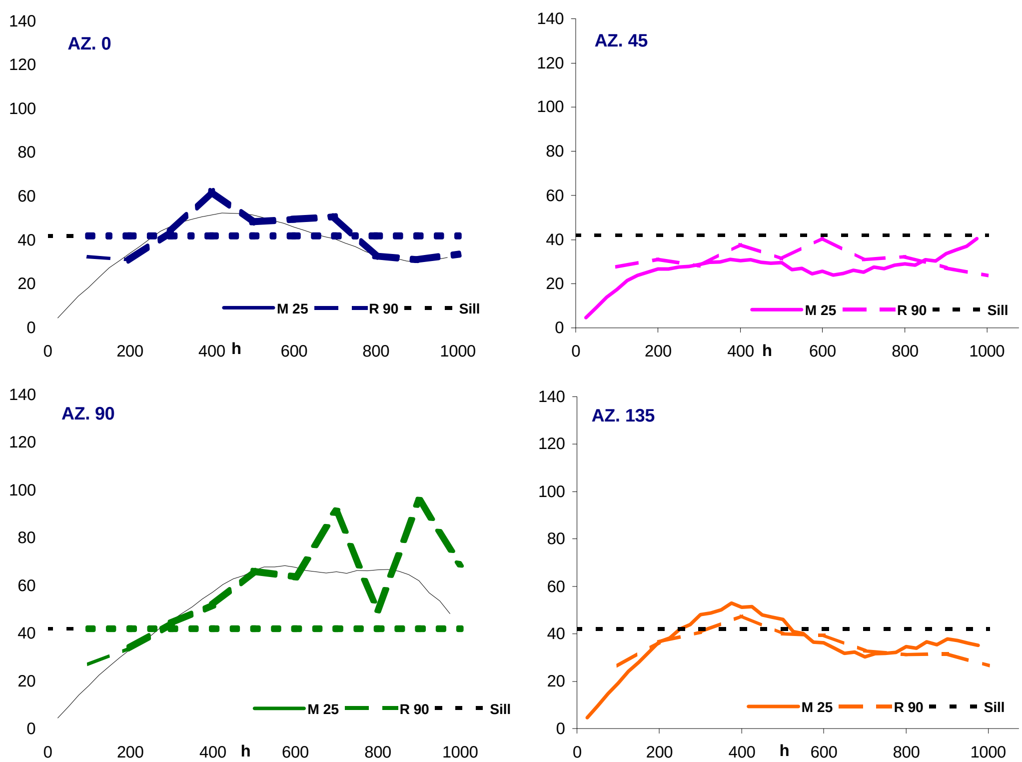

FIG-09 EXPERIMENTAL VARIOGRAMS R90 AND KTNVS VARIOGRAMS M25.

This set of variograms seem to fit much better than the ones in Fig-03. The strong hole effect in AZ 90-R90 looks more realistic with AZ 90-M25. Nugget effect seems more consistent and realistic in all of the M25 variograms.

However let us point out that KTNVS method is not fitting any model at all, nor it is doing any interpolation/smoothing to the initial variograms. KTNVS starts from the beginning, producing more data, then the problem due to lack of sampling has been overcome. Noise has been notoriously reduced, then structures are visible now and look consistent with the original set of experimental variograms. It is like: "Oh yes, now that I see this I would have set the variograms like that too".

Now we can test another match: the variograms-geology match to see how the two structures correlate each other. First lets take a look of a computer picture of the topography of the deposit, below in Fig-10.

Fig-10 below shows the topography of the deposit, to the left a perspective and to the right a plane view. Both figures showing the 90 bore holes.

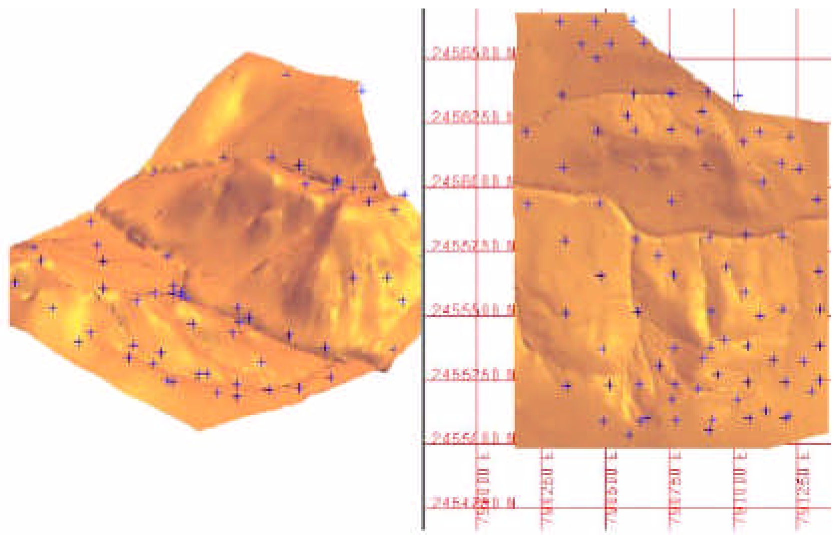

FIG-10 TO THE RIGHT A PERSPECTIVE VIEW OF THE CAO DEPOSIT THEN A PLANE VIEW, BOTH SHOWING THE 90 BORE HOLES.

Figure 11, below is the same deposit as Fig-10, but colored by ranges of CaO

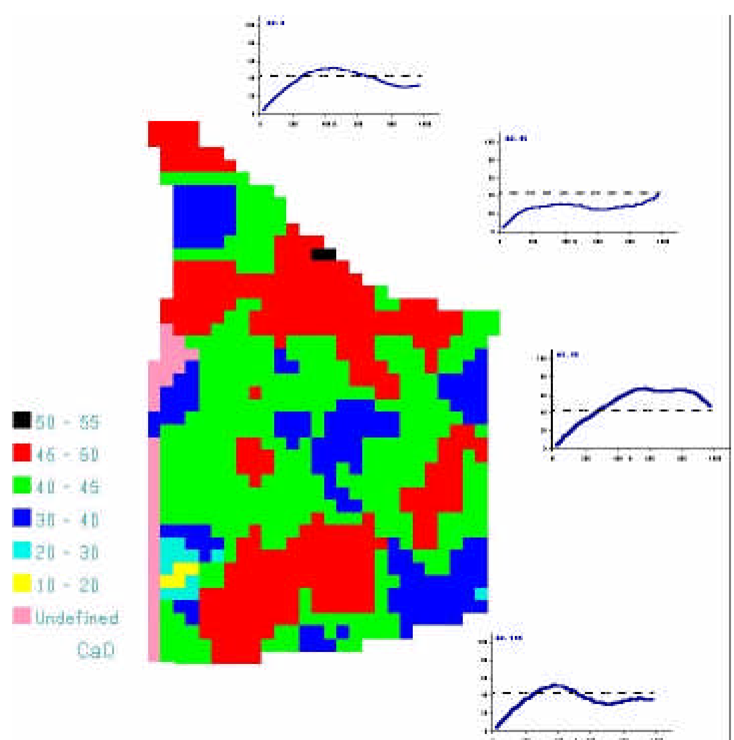

FIG-11 DEPOSIT WITH VALUES COLORED BY RANGES AND KTNVS VARIOGRAMS TO THE RIGHT.

Even though the horizontal scale in the variograms does not correspond to the horizontal scale in the deposit in Fig-11 above, there is however some visible correlation in the not so strong hole effect in AZ90 with geology in that direction. The same applies in AZ135 in its direction. It is also possible to find repeating events, which is the light hole effect of AZ90 and AZ135. Even the AZ 0 variogram gives us a close picture of a slight repetition too.

Now in the figure below we can see the same block picture but not the same kind of matching. The R90 variograms to the right, perhaps they are mismatching slightly different from geology.

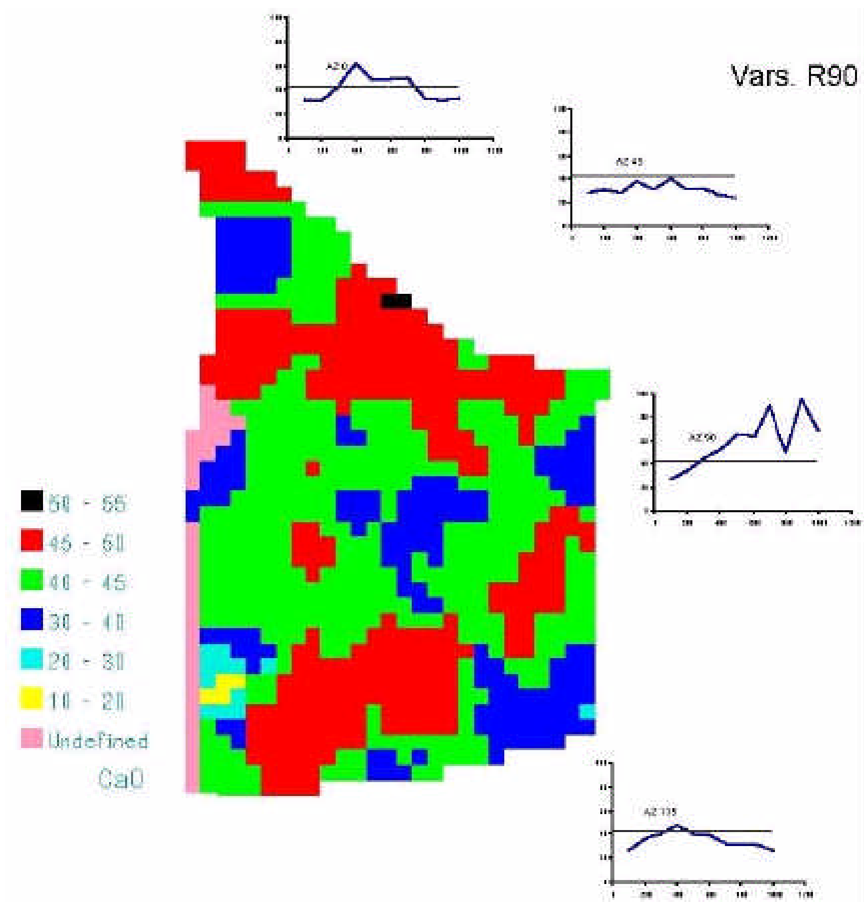

FIG-12 DEPOSIT AND VARIOGRAMS R90.

The reader can make its own comments comparing both figures: 11 and 12.

So far, perhaps we have eliminated noise and obtained clearer structures with KTNVS variograms. Perhaps we also find a correlation of structures between variograms and geology. Now we need the second part, which is to evaluate $\gamma(\boldsymbol{h})$ to calculate the entries of the correlation matrix.

V. The KTNVS 3D Autocorrelation Function

Next 3D figure is obtained from the KTNVS approach, shown below in Fig-13.

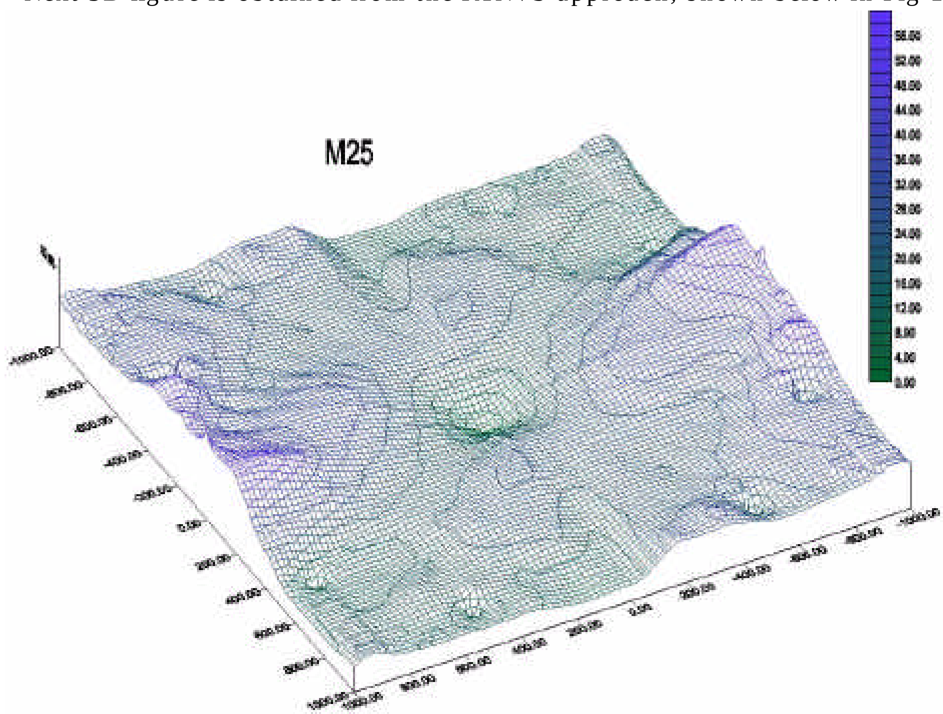

FIG-13 3D AUTOCORRELATION SURFACE FROM M25.

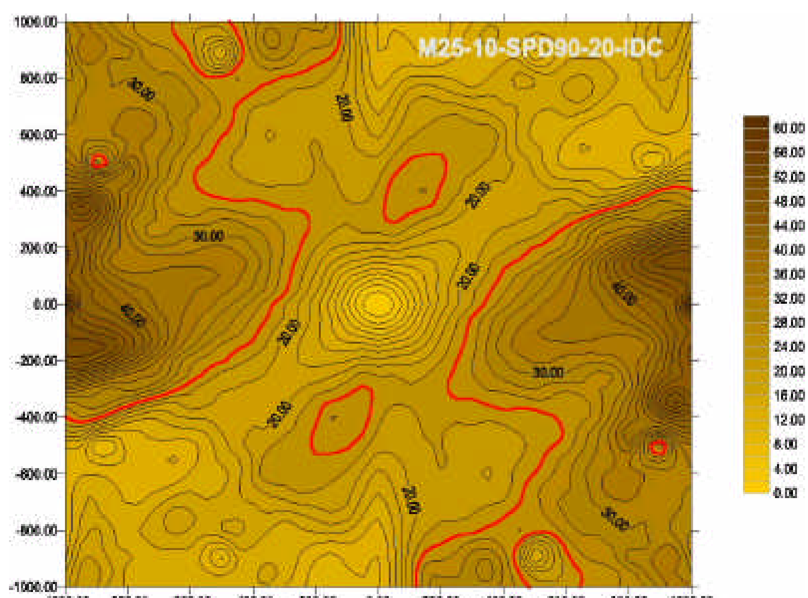

FIG-14 CONTOUR MAP FOR THE SURFACE IN FIGURE ABOVE.

Mapping the surface from the model(s), in the traditional way, now it is substituted by mapping directly the KTNVS surface. Although this time we do not have an equation to evaluate; what we have instead is a 3D empirical autocorrelation function to interpolate values from. The procedure is the same, not so the source of the evaluations.

From the KTNVS approach we can see, for instance, that there is a (slight) hole model, when and where is needed. Simple spherical alike models, cases of anisotropy/isotropy, a changing area of influence, (see Fig-14) not limited only to elliptic shapes, perhaps even more complex yet, perhaps not. Sophisticated models like cubic, gaussian, nested structures, and several others maybe not detected or found yet. All of them are implicit in the KTNVS surface, in a such detail as no theoretical model can do by now. Both surfaces in figures 04 and 13 are copied below, for a closer look.

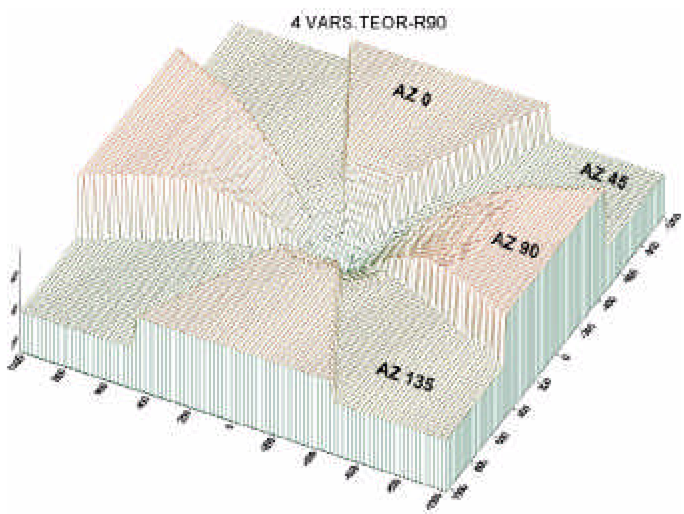

FIG-04 VARIOGRAM SURFACE FROM 4 THEORETICAL VARIOGRAMS.

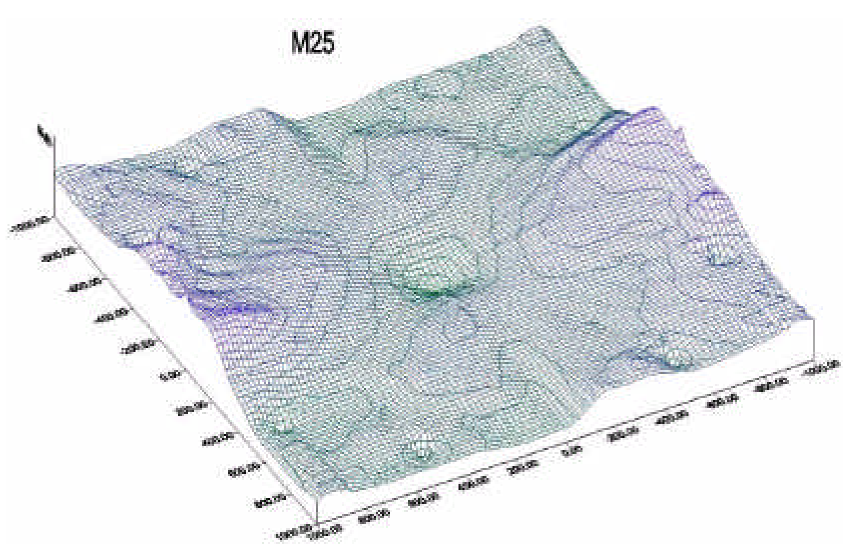

Fig-13 3D Autocorrelation Surface From M25.

Mapping KTNVS surface is a rather logical approach. Lets take a closer look of figure 13 above and show a close-up of it in figure 15 below:

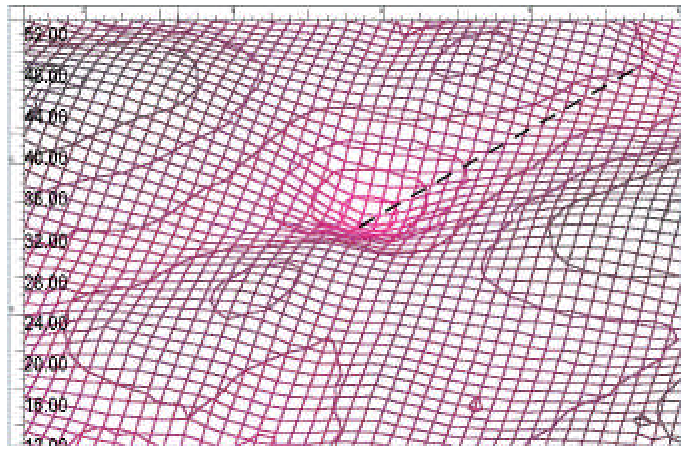

FIG-15 CLOSE-UP OF FIG-14. KTNVS SURFACE.

Let us say we select a couple of samples, they find in the direction shown in the figure, then depending on ***h***, and using interpolation, we do the mapping as it is shown in figure below.

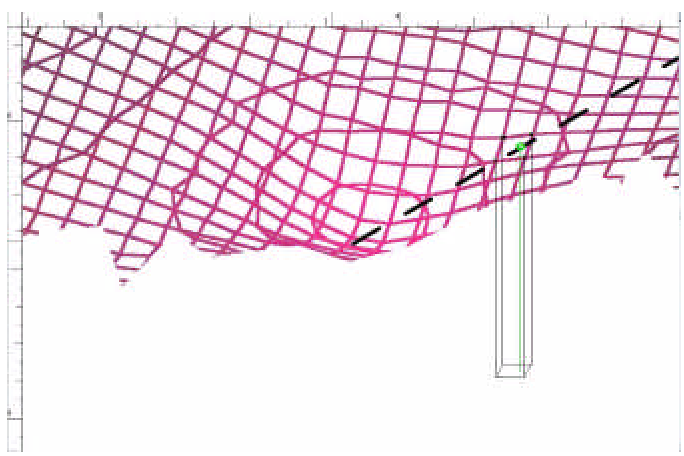

FIG-16 CLOSE-UP SHOWING INTERPOLATION IN A GIVEN DIRECTION.

Selecting couple by couple this way, without restriction in regard direction and no discontinuities, is possible to quantify $\gamma(\boldsymbol{h})$ without mathematical modeling .

Now, in the case of interpolations done in the traditional approach, for instance, using a surface like in figure 04. The interpolation seems to have some problems in the boundary of two adjacent models, when one sample is under a model and the second one lies below another different model.

RESULTS

VI. Case Study

The sampling used comes from a sedimentary CaO deposit also with values of SiO_2, Al_2O_3 and MgO.Only the CaO values was used. The size of the deposit is about 2025m north, 1175m east. The deposit is found on a slight creek. Shown in a figure the topography, magnified about 5 times from realty for the sake of the view, as well as a flat projection too. Small crosses show 90 bore holes.(see Fig-10)

The deposit was initially drilled with 60 holes (R60) 814 samples, then there was a second campaign with 30 more holes and 200 samples, to give a total of 90 holes (R90) and 1014 samples. In another study it was shown that the 200 late samples did not add much information to the initial database, one reason was that the size of the drilling was not the same as in the initial campaign so the variance was increased and the statistical behavior of this last sampling was rather close to be considered a different population sampling.

VII. Comparing Variograms

Variograms from the total drilling (R90) do not differ much from the first one (R60). See both sets of variograms below in figure 17.

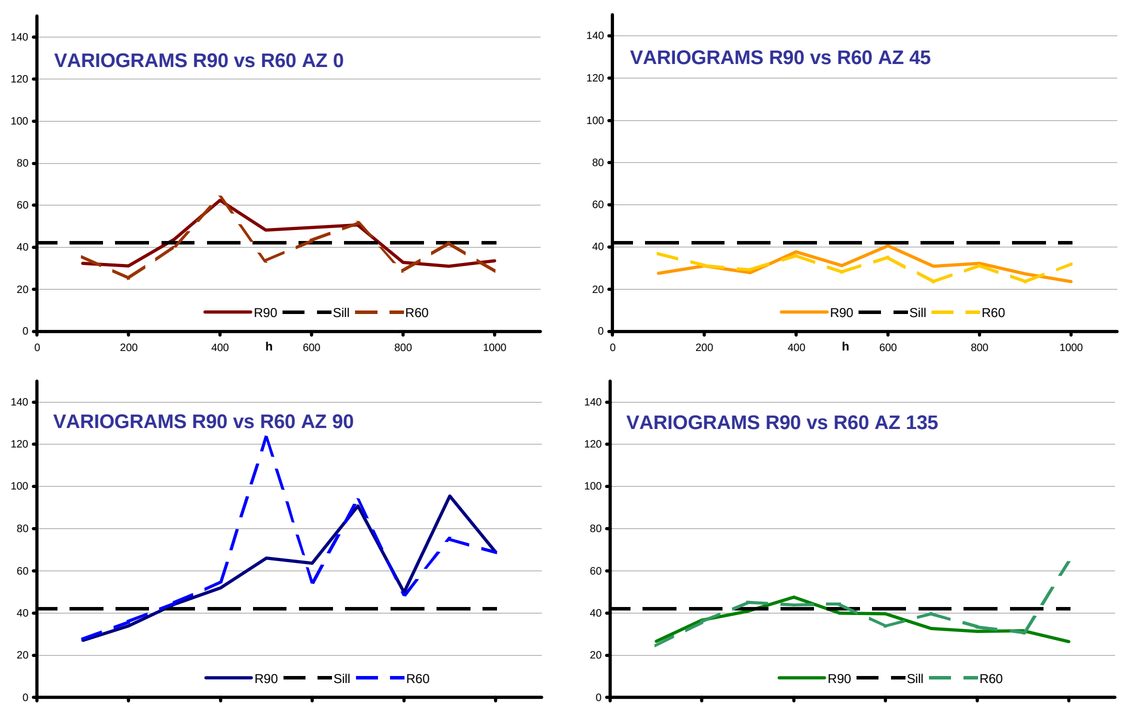

FIG-17 COMPARISON BETWEEN R90 AND R60 VARIOGRAMS, PAIR WISE SAME DIRECTION

There is not much difference between both sets of variograms, although in AZ90 R60 shows a stronger hole effect than R90. AZ45 keeps showing a pure nugget effect, maybe AZ0 too. However, some traces of spherical alike in AZ135.

Now matching R60 variograms and KTNVS variograms is possible to see there is also a good match, no matter if it was first or second campaign. This fact demonstrate two things: 1. The second drilling campaign was not necessary, and 2. KTNVS is not depending on the second or even more drilling campaigns. It is consistent to all of them. Compare figure below with figure 09

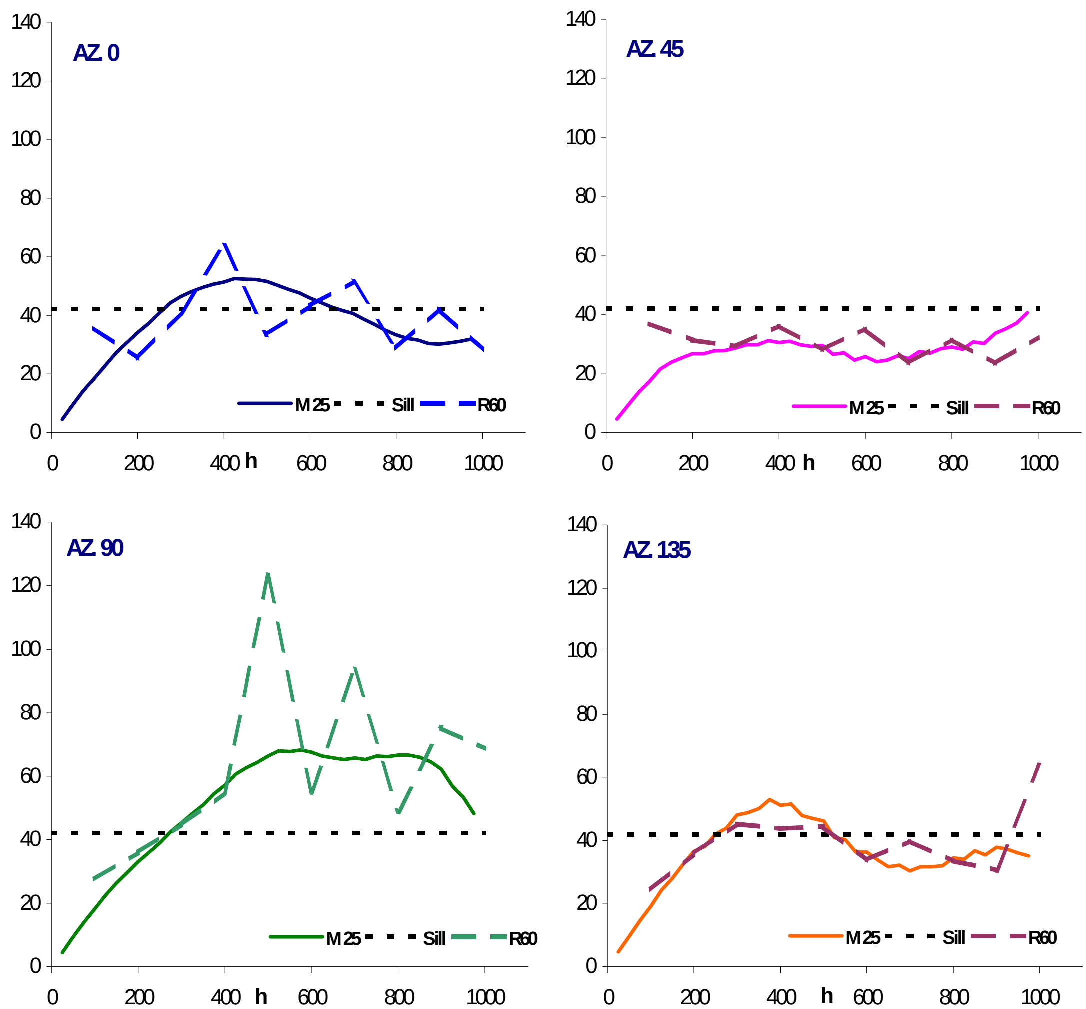

FIG-18 MATCHING EXPERIMENTAL VARIOGRAMS R60 AND KTNVS VARIOGRAMS.

Something else to be compared is both variogram surfaces: R60 and R90. The two figures below show surface and contour map, both from the experimental sampling R60, the first drilling campaign.

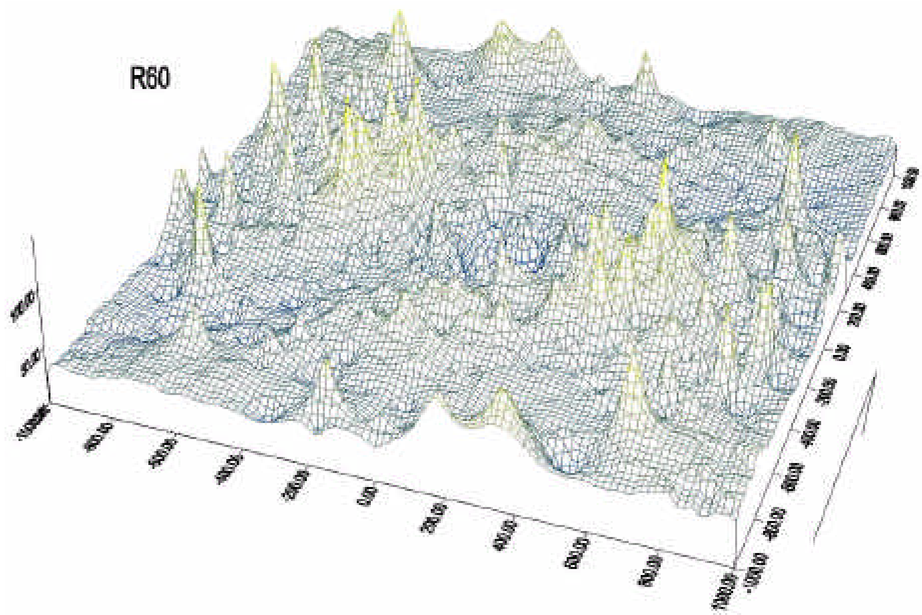

FIG-19 SURFACE VARIOGRAM FROM R60, FIRST DRILLING CAMPAIGN.

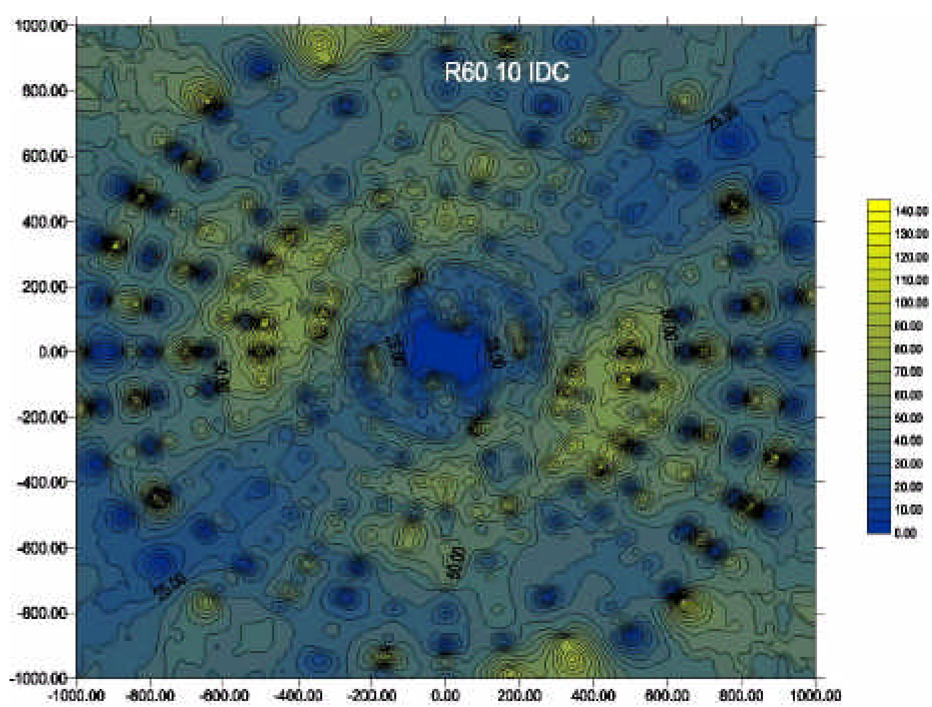

FIG-20 CONTOUR MAP FROM FIGURE 19 ABOVE.

Let us compare both figures above with the ones calculated from the total campaign, which is R90, in figures 06 and 07

VIII. Results from crossvalidation

The testing procedure used was crossvalidation, it refers to a test performed in the sampling itself. Eliminating a sample at a time and estimate it from the rest of them. Doing point kriging in this case. The advantage: it is possible to measure the error of estimate and no need to wait for the mine to be totally exploited.

To do this exercise, a set of 160 samples were chosen at random from the data set of 1014 samples. Same conditions were used both for traditional and KTNVS approaches. Same neighborhoods with the same number of samples, same size of equations and incognits. The only difference was the covariance matrix calculated first with traditional spherical models (fig-03). Then with the KTNVS method. In figure 20 below we have 3 parts: 1. The distribution of 160 CaO samples, selected at random from the total data. 2. Shows the distribution of the estimate using theoretical variograms and the traditional kriging, with spherical models. 3. Shows the KTNVS estimate distribution, kriging was done with a different covariance matrix, from the KTNV surface.

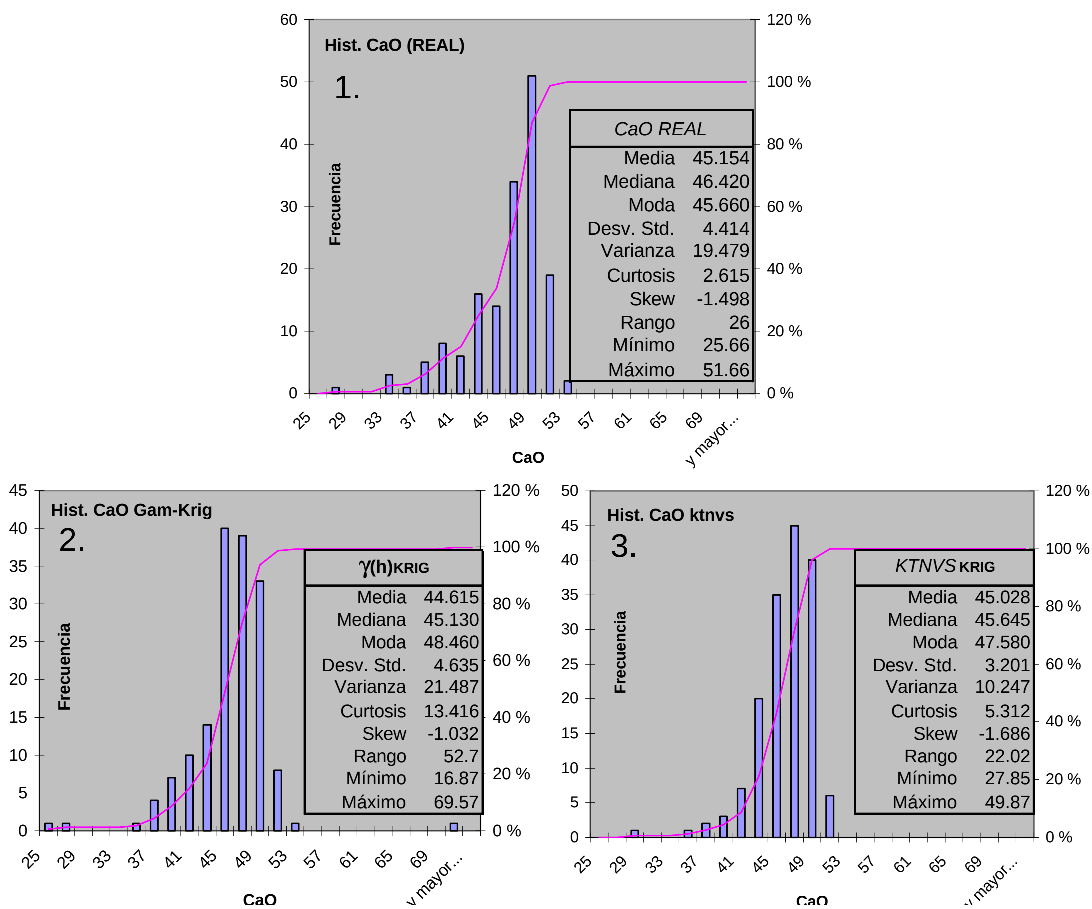

FIG-21 DISTRIBUTIONS AND BASIC STATISTICS OF: 1 ORIGINAL SAMPLES. 2 TRADITIONAL KRIGING, AND 3 KTNVS METHOD AND THEN KRIGING.

The following figure below is the estimation variance, and the error of estimate, statistics. The column to the left is the distribution of the estimation variance: Above, with the traditional models, below the KTNVS method. The column to the right shows the distribution of the error = real – estimate. Above the traditional and below the KTNVS.

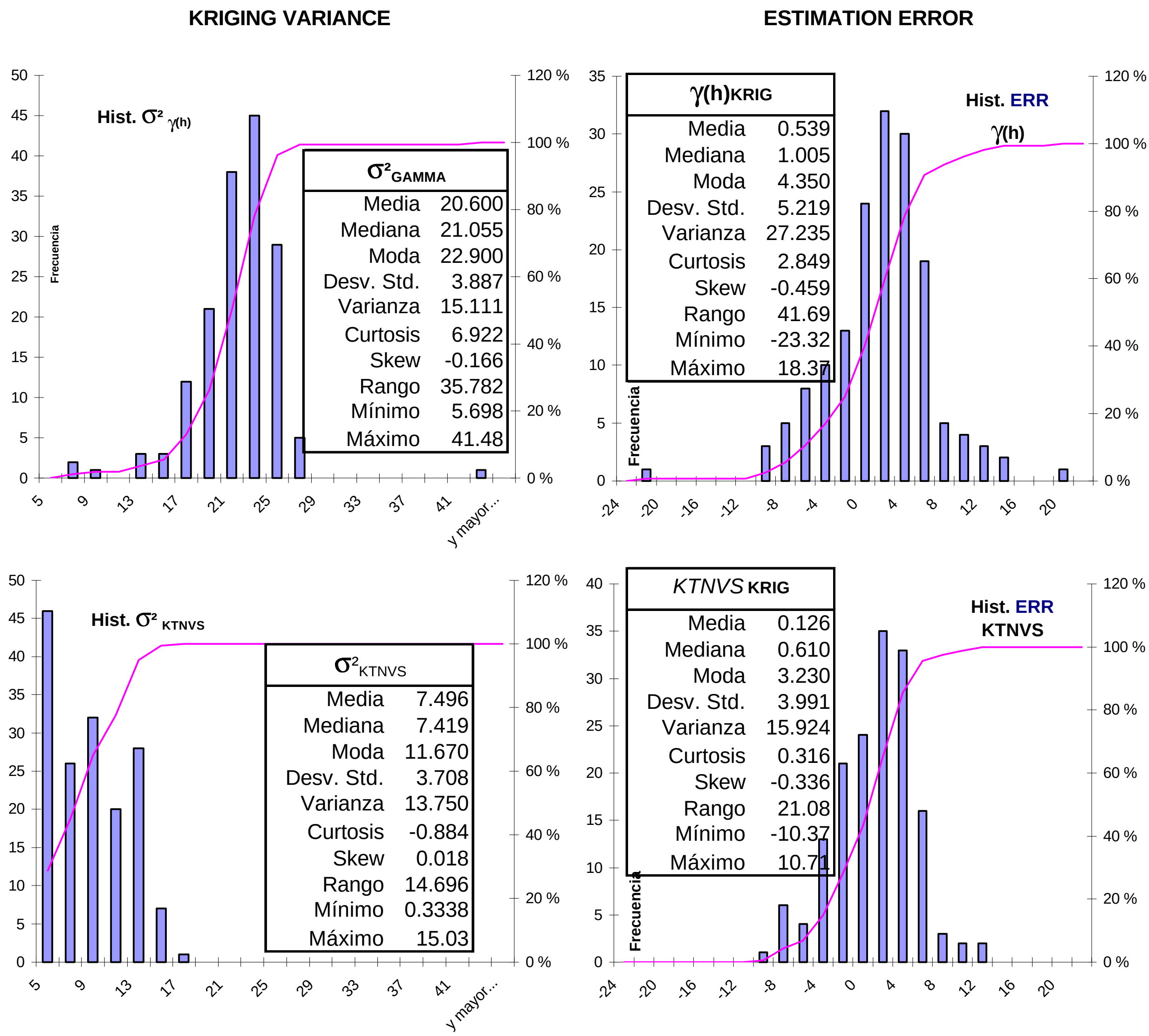

FIG-22 STATISTICS OF THE ESTIMATION VARIANCE AND THE ERROR OF ESTIMATE

Both figures to the left (above and below) show the basic statistics of the estimation variance: first the traditional then KTNVS. To the right (above and below) statistics of the estimation error. Not much to be added, perhaps both figures 21 and 22 speak for themselves.

CONCLUSIONS

The KTNVS method has demonstrated to be a rather natural and logic method that improves two important areas: the operational and costs. It presents a way to examine simulated sampling with fairly close realistic characteristics like real samples some times do. KTNVS can be helpful also in finding out benefits in augmenting sampling, (if that is the case) to evaluate, modify and study drilling patterns as many times as the user wants, without strong investments, done at computer speed. This is something impossible to do with real sampling; once it is done it is there for ever.

KTNVS has demonstrated with the case study presented here, that the investment of 30 more holes for the second campaign could have been saved. Then the conclusions are as follows:

1. When there is lack of sampling data and more drilling campaigns take time for a new budget to be approved, this method is a very fast and practical non-expensive way to enrich the database, compared to drilling costs.

2. It also gives a fast way for designing patterns for drilling, if this is the case, since it gives a more clear way to examine areas of influence and other geological structures.

3. Kriging estimation has been changed into a non-users dependant process since KTNVS gives a reliable and consistent method without the traditional variogram model-fitting.

4. It provides of clean variograms, (variogram surface) making it possible to have a better understanding of the geological structures of the deposit. Variography study is more effective now.

5. Computational processing without complex iterative algorithms that may end up with converging problems. It is a non-time consuming process.

6. This method reduces the error of estimate and kriging variance. The kriging estimation is improved.

Acknowledgments

The author wants to thank the Norwegian University of Science and Technology of Trondheim, and especially Professor Richard Sinding-Larsen for his important comments and advice, also to Karianne Enger for reviewing the manuscript, a person with different background may have encouraging opinions like a psychologist possible may.

APPENDIX

Performing a simulation where a flat deposit of 1500m by 1050m by 10m was designed and it was splinted in 17,750 blocks 10×10×10m size. In every single block we know the mean value in gr/TM of some kind of mineral.

Therefore the whole deposit is covered, since we know completely this particular population of blocks from our previous universe.

Then we know that the variograms from this population of blocks are the "true" variograms, there is no other source of random components but the intrinsic randomness of the deposit.

All the many random components affecting experimental variograms can be grouped in 4 basic categories: those who affect the value of the sample, those who affect the position of the sample, randomness itself of the deposit, and lack of data samples.

The exercise performed is a process of reducing the sampling in 8 stages, suppressing a percentage of the total of 17,750 samples each time. First take away 10% of the samples, then 20%, 30, 40, 50, 70, 90 and 92%. The last stage will leave only 1,260 samples. Each selection is a random draw, there is no other system on it. In every stage we calculate the four main variograms, then there is a set of 4 variograms for each stage. The variograms are shown below starting with the initial set, where there was no suppressing samples.

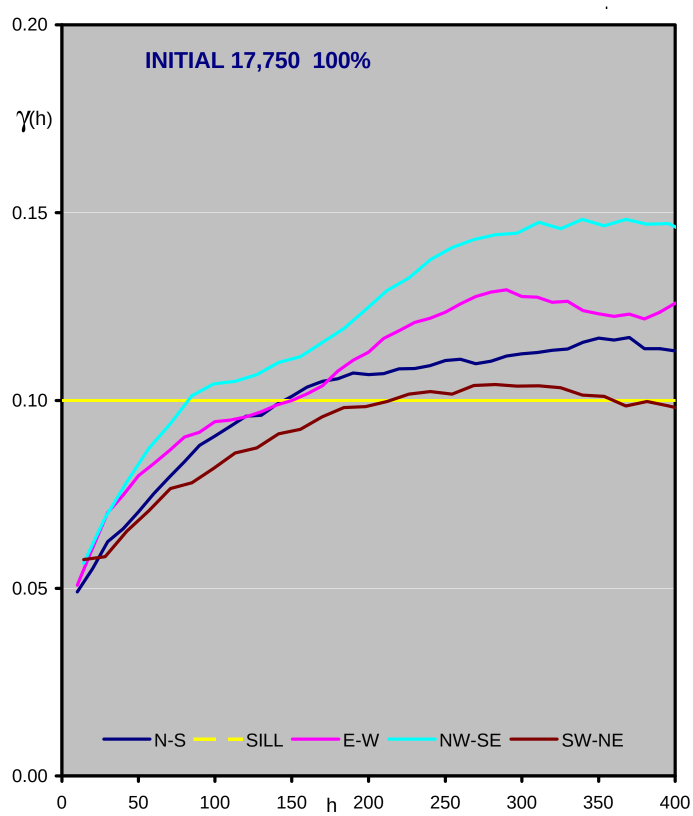

FIG-23 INITIAL SET OF VARIOGRAMS FROM THE POPULATION.

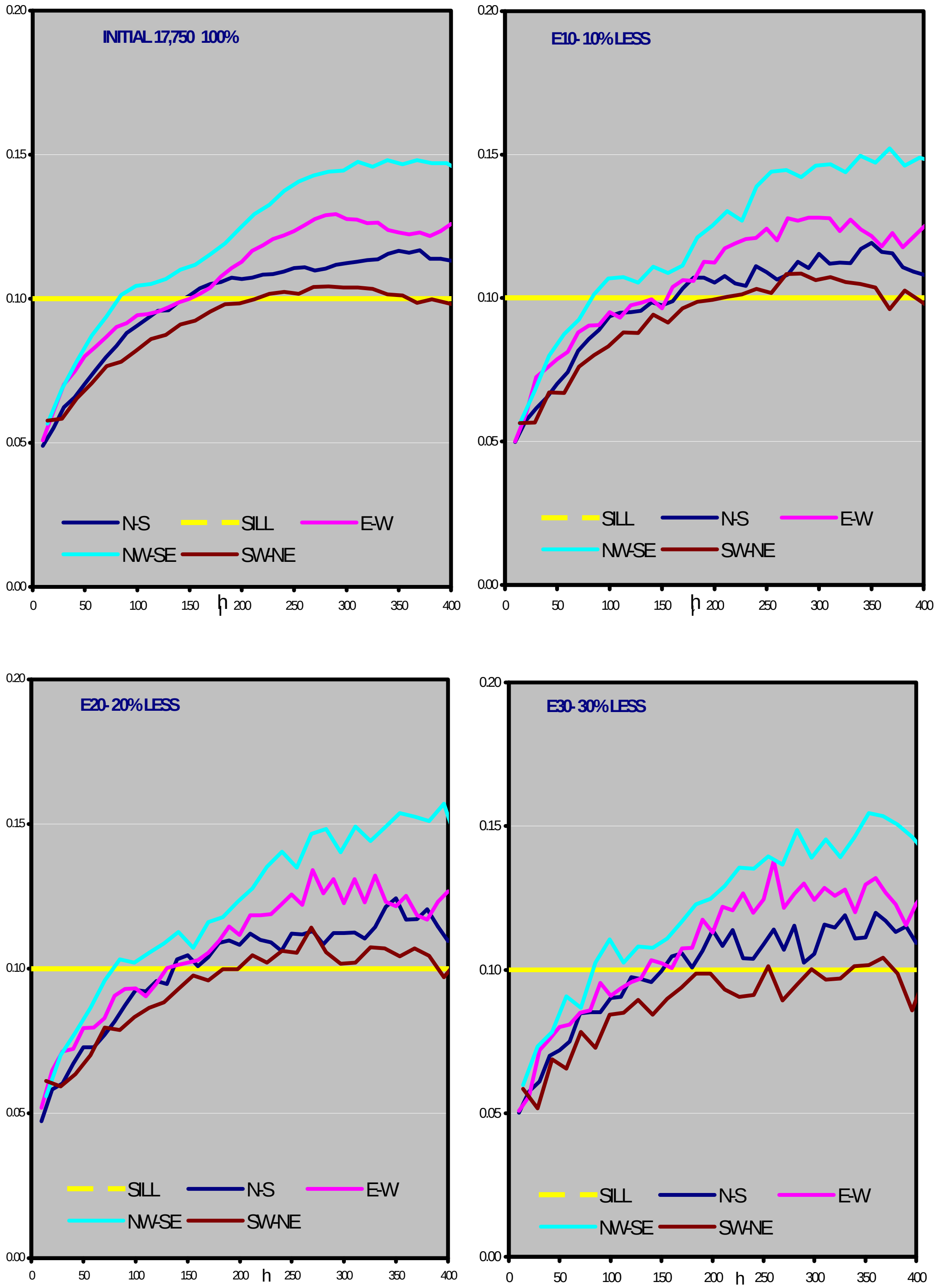

FIG-24 VARIOGRAMS SHOWING THE EFFECT WHEN REDUCING SAMPLES 10 TO 30 %.

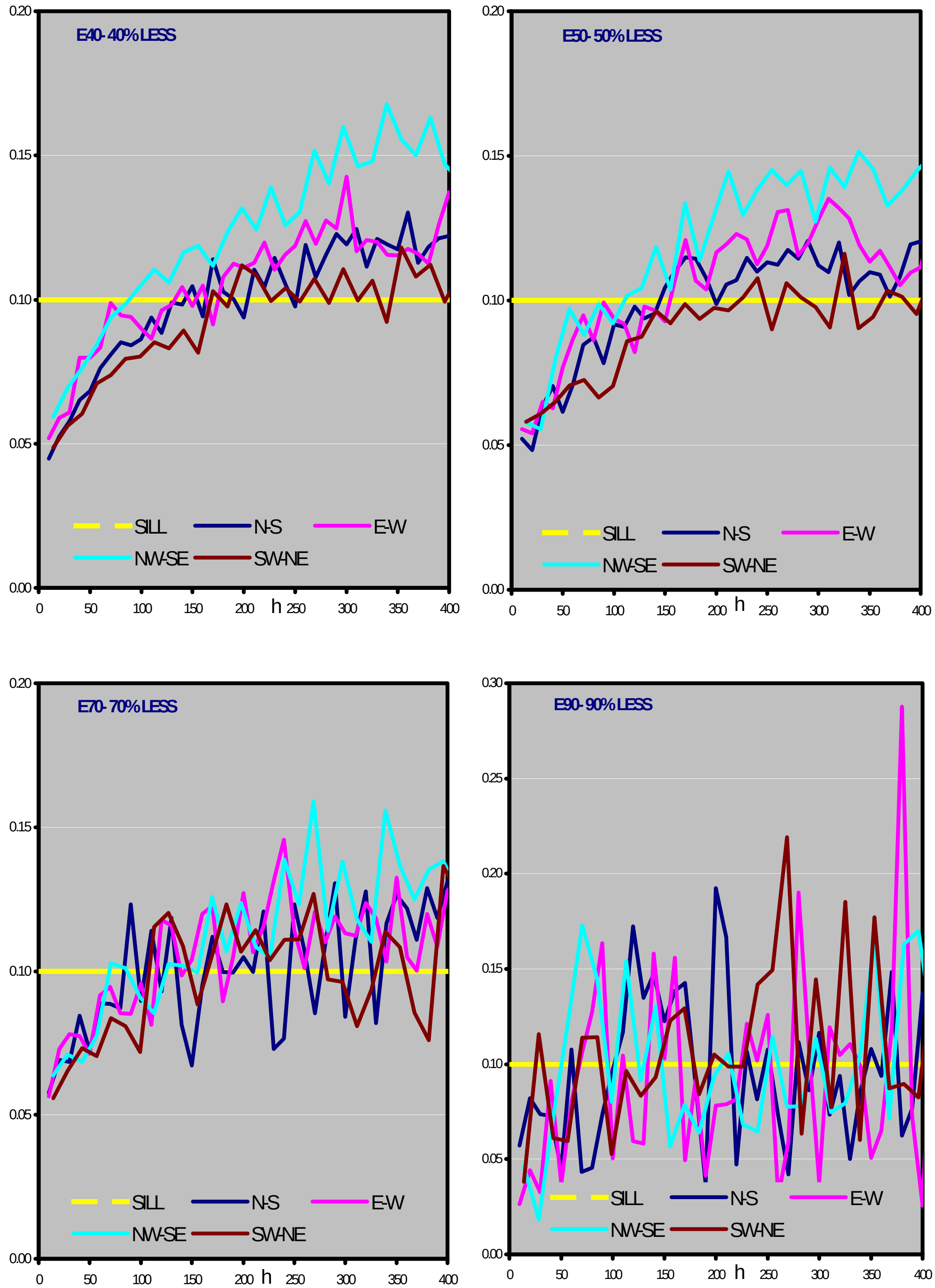

FIG-25 VARIOGRAMS SHOWING THE EFFECT WHEN REDUCING SAMPLES 40 TO 90 %.

Notice that the vertical scale has been reduced in plot for E90 and E92 (below) so we can see it complete. The following figure below shows 92% less samples which is close to the number of samples used in the case study, and we can see the variograms below are rather similar to the R60 / R90 variograms.

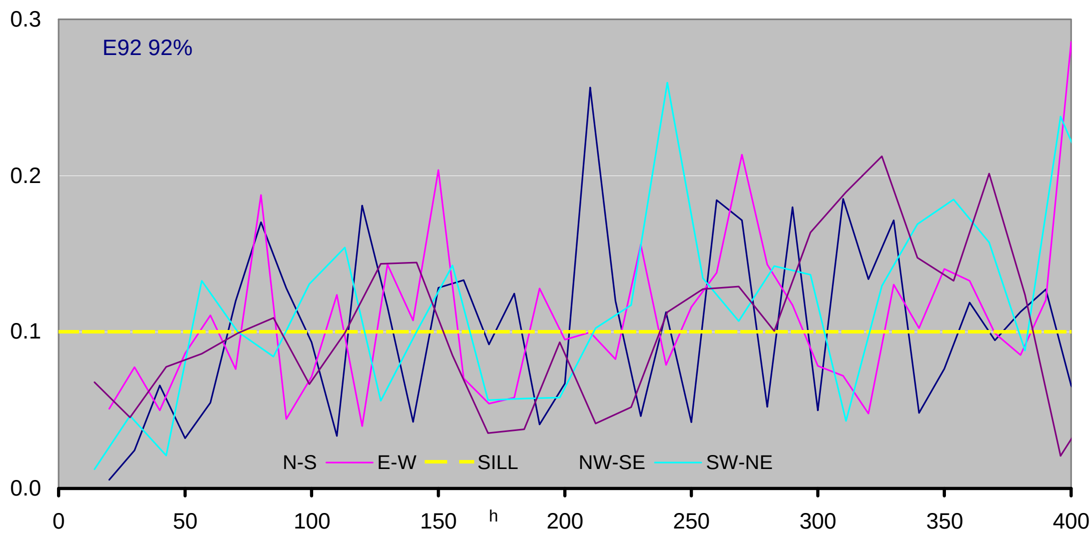

FIG-26 VARIOGRAMS WITH 92% LESS SAMPLES

Let us compare variograms in figure 26, with variograms in figure 02.

The following figure bellow is matching the initial variogram, no sampling reduction, and E92, in N-S direction. This is the basic principle of KTNVS. Few sampling *vs.* rich sampling.

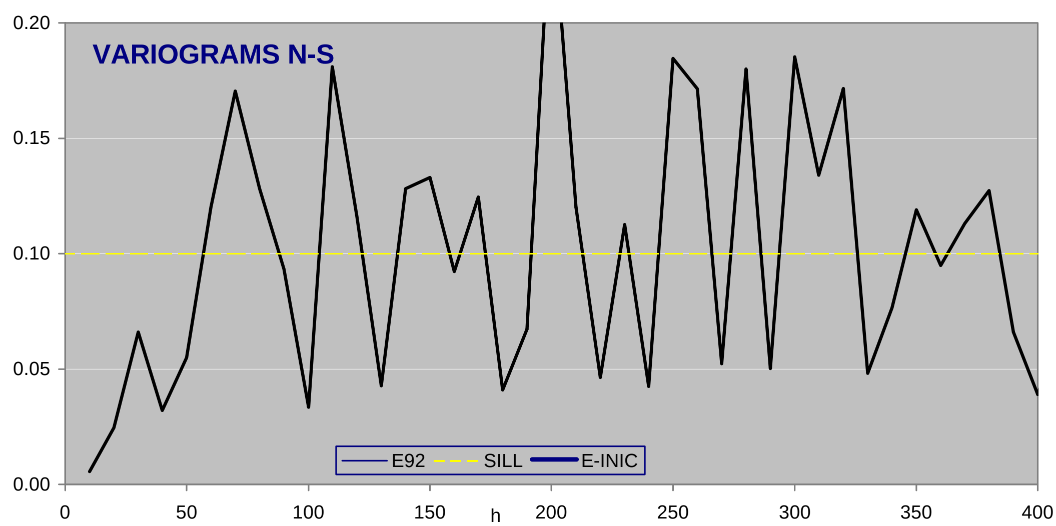

FIG-27 MATCHING VARIOGRAMS WITH 100% SAMPLES VS 92% LESS SAMPLES IN N-S.

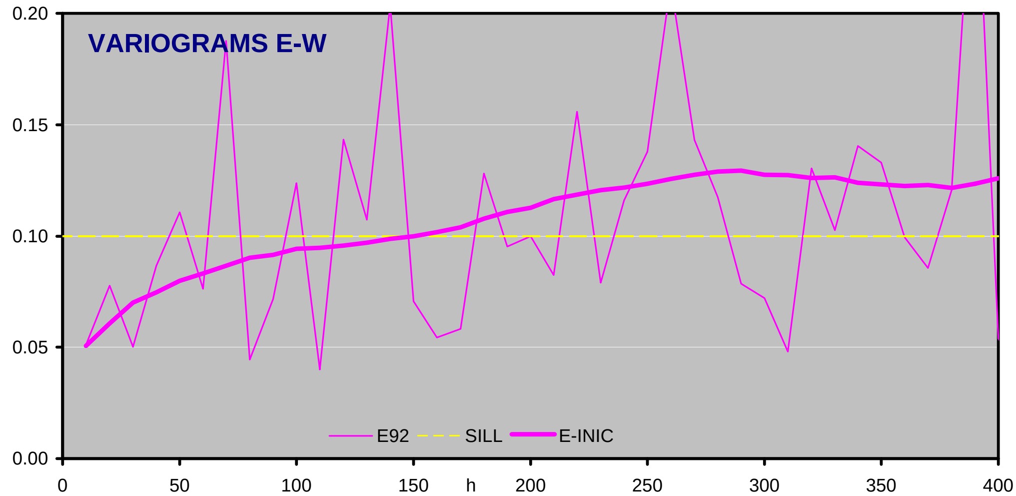

FIG-28 MATCHING VARIOGRAMS WITH 100% SAMPLES VS 92% LESS SAMPLES IN E-W.

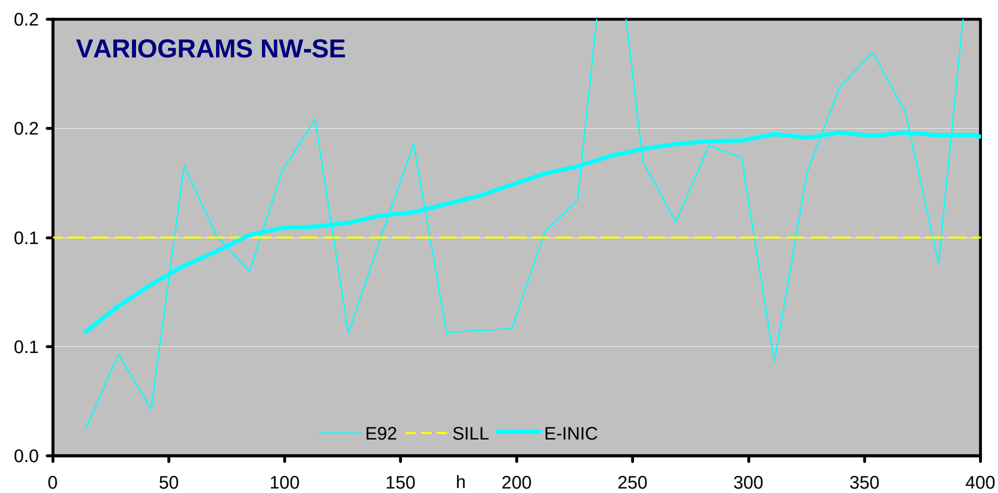

FIG-29 MATCHING VARIOGRAMS WITH 100% SAMPLES VS 92% LESS SAMPLES IN NW-SE.

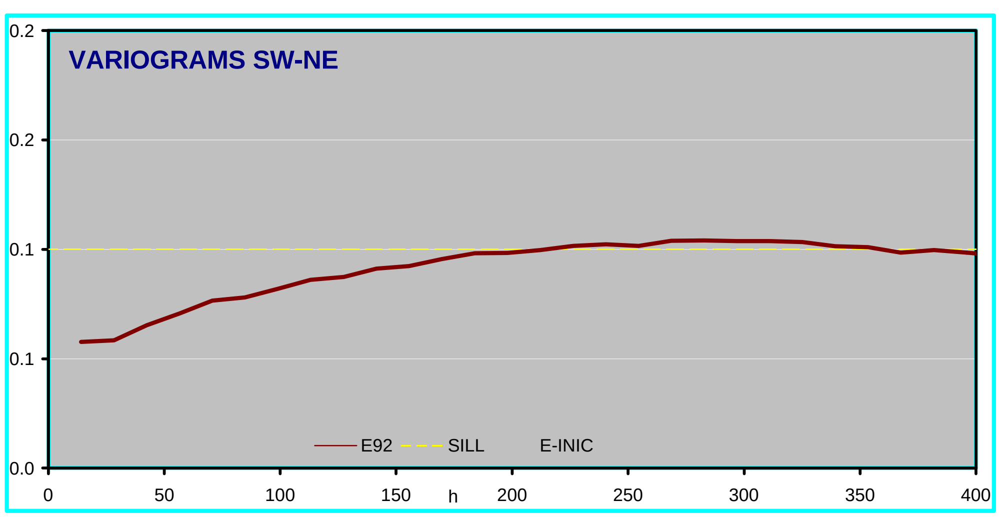

FIG-30 MATCHING VARIOGRAM WITH 100% SAMPLES VS 92% LESS SAMPLES IN SW-NE.

REFERENCES

Adisoma, G.S., Kim, Y.C., JACKKNIFE KRIGING – *a Simple and Robust New Estimator, Montreal, Canada. 1993.*

Armstrong, M. BASIC LINEAR GEOSTATISTICS: *Springer-Verlag Berlin Heidelberg New York. 1998.*

Clark, I., DOES GEOSTATISTICS WORK?, *University of Arizona, 1979*

Clark, I., PRACTICAL GEOSTATISTICS: *Applied Science Publishers Ltd., London, England. 1979.*

David, M. GEOSTATISTICAL ORE RESERVE ESTIMATION, *Elsevier Scientific Company, Amsterdam 1977.*

Goovaerts, P., GEOSTATISTICS FOR NATURAL RESOURCES EVALUATION: Oxford University Press. 1997.

Isaaks, E.H., Srivastava R.M., AN INTRODUCTION TO APPLIED GEOSTATISTICS, *Oxford University Press, New York, 1989.*

Mejía, E., ENTENDIENDO EL VARIOGRAMA, *AIMMGM, México, 1987.*

Mejía, E., EL EFECTO DEL EFECTO PEPITA, *AIMMGM, México, 1993.*

Mejía, E., SIMULATOR OF GEOLOGICAL SECTIONS, ORE DEPOSITS AND ITS VARIOGRAPHY, *APCOM. Montreal, Canada. 1993.*