Inference of the Nugget Effect and Variogram Range with Sample Compositing

David F. Machuca Mory and Clayton V. Deutsch

Centre for Computational Geostatistics
Department of Civil and Environmental Engineering
University of Alberta

The correct determination of the nugget effect is critical when fitting the variogram model for estimation, simulation and other geostatistical tasks. When exploration data are used, the vertical or down-the-hole direction provides sufficient data to determine this parameter; however in certain cases defining its correct value, even from these direction of highest data density, is not so obvious. The idea presented here is to use the variance of data composited to different scales. These experimental dispersion variances are very sensitive to the short scale variogram. An iterative process is implemented to find the variogram range that allows the convergence to a single value of the nugget effect. The algorithm proves useful to calculate the nugget effect of synthetic and real data when their correspondent variograms are noisy.

Introduction

The nugget effect can be understood as the grade variability at a resolution or scale shorter than the sample size and/or sampling separation and appears as a discontinuity of the variogram near the origin. The presence of a component of the mineralization with a range of continuity shorter than the sample support, the short scale spatial variability that can not be not discerned with the given sample separation, and measurement or positioning errors, have been indicated as the sources of nugget effect (Chilés and Delfiner, 1999).

In Kriging, a high nugget effect leads to smoother estimates except near data locations, where artifact discontinuities become more pronounced (Goovaerts, 1997). In Geostatistical simulation, resulting realizations show less structure and increased randomness, that is translated to an increased smoothing of grades when point scale realizations are upscaled to the SMU's size. In the same way, in global change of support problems, the SMU scale distribution will become less variable and more symmetric as the nugget effect increases.

Therefore, the correct determination of the nugget effect value has a very important affect on the outcome of Geostatistical methods that require a variogram model. The nugget effect is considered isotropic, thus its value is commonly determined from the vertical or down-the-hole variograms, since sampling separation is the closest in these directions, and then applied to the variogram model fit in all directions. However, in some cases, choosing the correct nugget effect for variogram modelling can be difficult because of noise in the data, software limitations or geomodeler inexperience.

The idea of this paper is to use the composite samples variance to infer the correct nugget effect using the equation for the variance of the composite values. However, if the variogram range is unknown, a simple optimization algorithm can be applied to find the variogram range that allows the convergence of the nugget effect values calculated from the composite grade variance at different lengths. Thus we obtain both the nugget effect and a first estimate of the variogram range from composite variances.

Theoretical Background

If the original sampling intervals are of constant length, the variance of drillhole sample composites at a given length L can be calculated by (Isaaks and Srivastava, 1989):

$$
\begin{aligned}
Var\{Z_L\} = \sigma_L^2 &= \frac{1}{n_L^2}\sum_{i=1}^{n_L}\sum_{j=1}^{n_L} Cov\{z_i, z_j\} \\
&= \frac{\sigma^2}{n_L} + \frac{1}{n_L^2}\sum_{i}^{n_L}\sum_{j\neq i}^{n_L} Cov\{z_i, z_j\}
\end{aligned} \tag{1}
$$

Where σ_L^2 is the composite variance for a composite length equal to *L*, n_L is the number of original samples per composite, and σ^2 is the variance in the original sampling scale. The covariances between the samples within the composites, $Cov\{z_i, z_j\}$, can be rewritten as:

$$
\mathrm{Cov}(z_i, z_j) = \sigma^2 - C_0 - C_1 \cdot \gamma(z_i, z_j) \tag{2}
$$

Where C_0 is the point scale nugget effect, C_1 is the variogram differential sill and $\gamma(z_i, z_j)$ is the standardized variogram between samples *i* and *j* within the composite. Thus, expression (1) can be approached by:

$$
\begin{aligned}
\sigma_L^2 &\approx \frac{\sigma^2}{n_L} + \frac{n_L^2 - n_L}{n_L^2}\left(\sigma^2 - C_0 - C_1 \cdot \gamma(\bar{h})\right) \\
&= \sigma^2 - (C_0 + C_1 \cdot \gamma(\bar{h}))\left(1 - \frac{1}{n_L}\right)
\end{aligned} \tag{3}
$$

Where $\bar{h}$ is the average separation between original samples within a composite, and $\gamma(\bar{h})$ is the variogram value at the later distance separation. Then, the point scale nugget effect can be calculated by inversing the previous equation, as:

$$
C_0 \approx \frac{\sigma^2 - \sigma_L^2}{\left(1 - \frac{1}{n_L}\right)} - C_1 \cdot \gamma(\bar{h}) \tag{4}
$$

Since $C_1 = \sigma^2 - C_0$, expression (5) becomes:

$$
C_0 \approx \frac{1}{C(\bar{h})} \cdot \left(\frac{\sigma^2 - \sigma_L^2}{\left(1 - \frac{1}{n_L}\right)} - \sigma^2 \cdot \gamma(\bar{h}) \right) \tag{5}
$$

The correct point scale nugget effect value derived from this expression should be constant whatever the composite length. However, since the variogram model can be also difficult to fit if the down-the-hole variogram is very discontinuous, the correct model type and range must be found. In order to find this correct down-the-hole variogram, the proposed idea is to test recursively different model ranges until the one that fulfills the criterion is found. This can be accomplished using a simple optimization algorithm that minimizes the absolute value of the slope of the Nugget Effect values calculated with different composite lengths and variogram ranges. The implementation of this optimization procedure is described next.

Algorithm Description and Implementation

Given a set of drillhole interval samples, the program `nuggcalc` implemented in FORTRAN calculates nugget effect for the original the sample scale and the down-the-hole, sample scale, variogram range for a chosen variogram model. The procedure implemented in this program can be summarized in the following steps:

1. Read a given drillhole data file and calculate the variance of original samples.
2. Calculate the composite values and their variances at different composite lengths, these lengths are entire multiples of the original sample length.
3. Calculate expression (5) for a range of composite sizes with a given variogram range guess, a_0.
4. Fit a regression line to the nugget effect values obtained for different composite sizes, keep the line slope, s_i, and its y-axis intercept, $C_{0(i)}$.
5. Slightly perturb the previous variogram range and calculate again expression (5) for the same range of composites sizes using the new variogram range a_{i+1}
6. Fit a new regression line to the nugget effect values obtained for different composite sizes and keep the new slope, s_{i+1}, and intercept, $C_{0(i+1)}$.
7. If $|s_{i+1}| \le |s_i|$ keep the new variogram range and intercept, else, retrieve the previous values.
8. Iterate several thousand times from point 5.
9. After *n* iterations the final y-axis intercept, $C_{0(n)}$, and range a_n are kept as the point scale nugget effect and variogram range, respectively.

Additionally to the estimated nugget effect and down-the-hole variogram range, this program provides a composite samples file at a chosen length.

The parameter file for this program is presented in the next figure:

```
                Parameters for NUGGCALC
                ***********************

START OF PARAMETERS:
ddh.dat                        - data file with assays
1 2 3 4 5 6 7 8                - DDHID, x_mid, y_mid, z_mid, from, to, lenght, grade
-1.0e21     1.0e21             - trimming limits
nuggcalc.sum                   - file for summary output
comps-4.5m.out                 - file for composited samples
1.5                            - Original sample lenght
2  8                           - Compositing multiples
3                              - Multiple number for composites file
40                             - variogram range guess
1                              - variogram type 1- spherical, 2-exponential
2000                           - number of iterations
69069                          - random number seed
```

Figure 1: Parameters for the program `Nuggcalc.`

The application of this algorithm is illustrated in the next examples using synthetic and real data.

Example 1: Synthetic Data

The synthetic data used to testing the program in this first example consists of a single hypothetical drillhole with 30,000 samples at intervals of 1m long containing the values of 5 gaussian distributed variables generated using sequential Gaussian simulation with a spherical variogram of 50m range but with 5 different nugget effect values.

In a first step the correct known variogram range was provided as an input to the program, thus, the optimization of the range was not necessary. Figure 2 show that if the true variogram range is known, the inferred nugget effect is kept effect is almost constant and very close to the correct nugget effect of simulated data, for any composite length used.

Conversely, if an incorrect variogram range is used, the nugget effect values calculated considering different composite lengths follow lines with positive slope if the variogram range guess is higher than the true variogram range, and with negative slope in the contrary case (See figure 3).

Figure 4 shows the evolution of the slope of inferred nugget effect values according to the variogram range used for two different simulated variables, which were generated using a nugget effect of 0.2 and 0.4, respectively. In this figure it can be observed that underestimation of the variogram range has a bigger impact than overestimation in the slope of the estimated nugget effect values. Independently of the nugget effect, the curves intersect the zero line around 52m, which is concordant to the true variogram range for this data (50m).

Finally, figure 5 present the decreasing of composite values variance as the composite sample length increases.

Example 2: Disseminated Gold Deposit

For this first practical example the data set chosen consists of 111 vertical and sub-vertical drillholes from an exploration campaign in a disseminated gold deposit located in Central America. This data set contains 7156 gold grade assays in intervals of 1.5m length.

The down-the-hole experimental variogram for this data set is shown in the figure 6. As it can be observed in this figure, fitting the correct nugget effect and variogram range is not so simple in this case.

For this particular test, the program `nuggcalc` was run with a limited number of different variogram range guesses without the optimization of this parameter. The range value guess that yields a linear regression fit with the closest to zero slope is 13m, the estimated nugget effect with this range is 0.46 (figure 7). Differently from the previous synthetic example, the estimated nugget effect values at different composite lengths do not form straight lines, but they show appreciable deviations from the trend line at small composite lengths (See figure 7).

These fluctuations in the estimated nugget effect are related to data induced composite variance fluctuations at short composite lengths that can be observed in the curve relating the variance of composite samples and the length of composites (figure 8).

Figure 9 show the curve of variogram range guesses vs. the slope of the estimated nugget effect fit, the intercept at the zero slope line is given around 13m, which is consistent with the range observed in the experimental down-the-hole variogram (Figure 6).

Figure 10 shows the experimental down-the-hole variogram and the spherical variogram model fitted with the nugget effect and variogram range values of 0.46 and 13m, respectively. As it can be observed there, the variogram model thus obtained is a fairly satisfactory fit to the experimental variogram.

Example 3: Copper/Zinc Massive Sulphide Deposit

The data set for this last example was taken from the companion CD of the book “Practical Geostatistics, Modeling and Spatial Analysis” (Houlding, 1999). This data correspond to a volcanogenic massive Cu and Zn sulphide deposit. Copper and Zinc grades are in 2m length intervals contained in the 16 closest to vertical drillholes. Cu grade nugget effect calculation was performed iteratively with a minimum composite length of 2m, and a maximum of 50m.

It is important to remark that although there exists in this data set drillholes with inclinations ranging from +27° to -88°, only the closest to vertical drillholes were used, since the variogram range can vary considerably in other directions, and thus, invalidating the algorithm.

Figure 11 present the down-the-hole experimental variogram of Copper grades and the model fitted using the parameters obtained using the full optimization algorithm. The single structure variogram model obtained fits very well the experimental variogram. However, if the experimental variogram show breaks and other features that could be better modelled with several nested structures, the results obtained by single structure approach implemented in the algorithm can be misleading and only valid for the shortest distances. This can be appreciated in the figure 12, where the experimental down-the-hole variogram of Zinc grades is plotted against different variogram models with parameters obtained using different maximum composite lengths.

Note that in figure 12 that the maximum composite length was allowed to exceed the variogram range, otherwise the calibrated nugget effect and variogram range is representative only for lag distances inferior to the maximum composite length.

Discussion and Conclusions

The variances of data composited to different lengths are used to provide estimates of the nugget effect and variogram range. A program is implemented for this purpose. The iterative procedure is straightforward to implement and to apply. The algorithm will work best for composite lengths that are relatively short, for single-structure variograms and for variograms without cyclicity. Moreover, the approach requires data of constant input length. These limitations could be overcome in a more sophisticated program.

The short scale behaviour of the variogram is of paramount importance. The resulting nugget effect and estimate of the variogram range are very useful to the practitioner who has to fit many variograms and wants the short scale behaviour to be consistent with the data.

References

Chilès J.P. & Delfiner, P., 1999. *Geostatistics: modeling spatial uncertainty*. Wiley-Interscience, New York.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.

Houlding, S.W., 1999. *Practical Geostatistics, Modeling and Spatial Analysis*. Springer-Verlag, Berlin Heidelberg.

Isaaks, E. H. and Srivastava R. M., 1989. *An introduction to Applied Geostatistics*. Oxford University Press, New York.

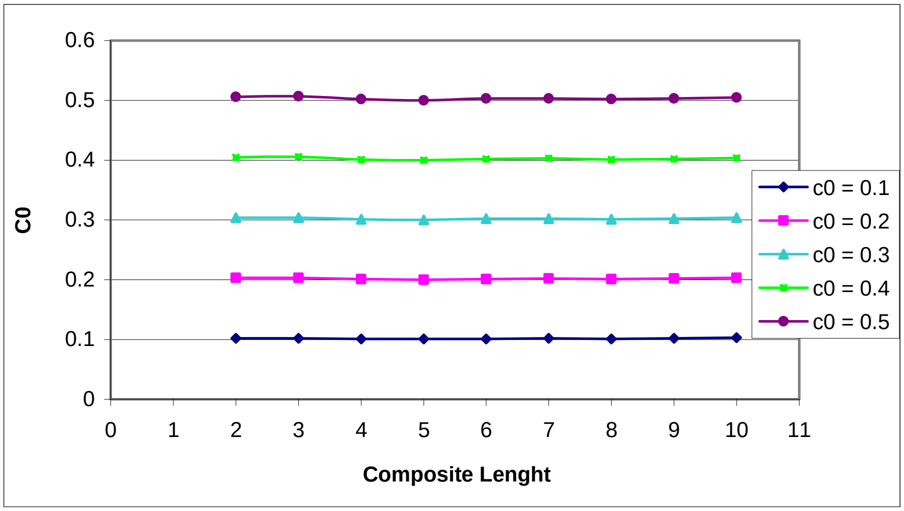

Figure 2: Calibrated point scale nugget effect using different composite lengths for five different variables simulated with different nugget effect values. The original sample length is 1 unit and the true variogram range is known.

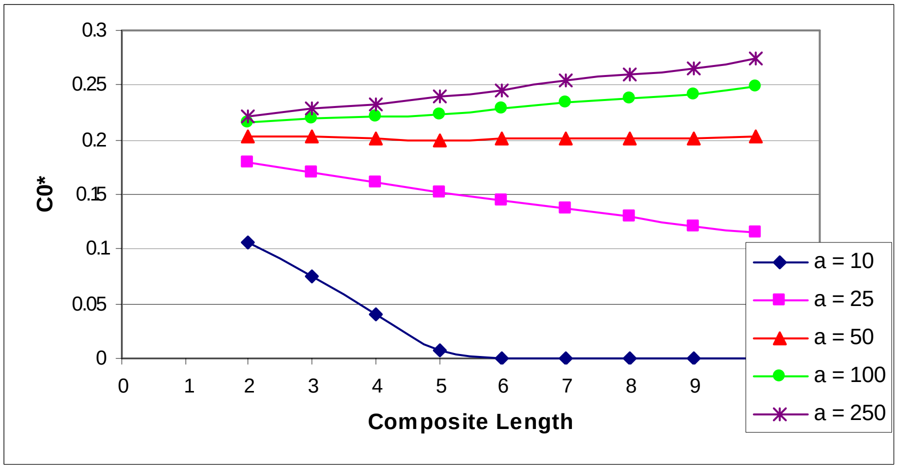

Figure 3: Calibrated nugget effect using different composite lengths and several variogram range guesses (10, 25, 50, 100, and 250 units) for a variable simulated with a nugget effect of 0.2 and a spherical variogram range of 1.

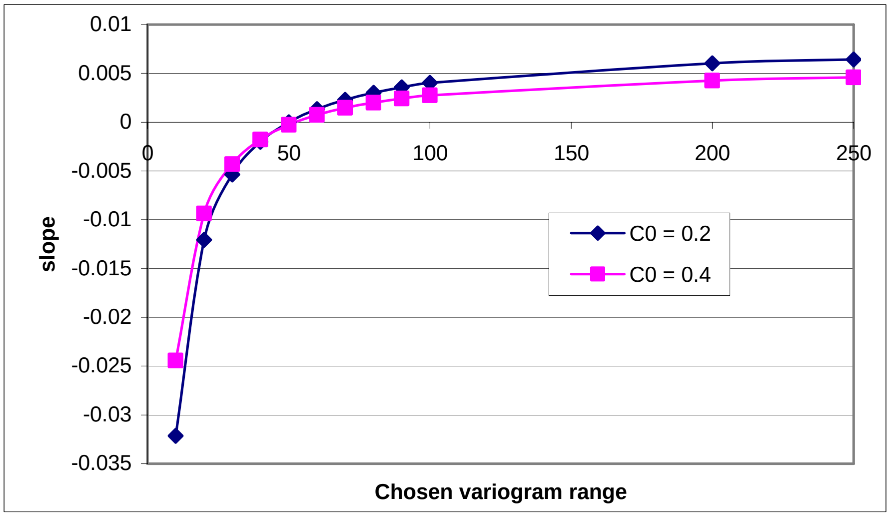

Figure 4: Change in the slope of the inferred nugget effect vs. composite length regression line, at different variogram ranges and for two variables simulated with nugget effect values of 0.2 and 0.4, respectively.

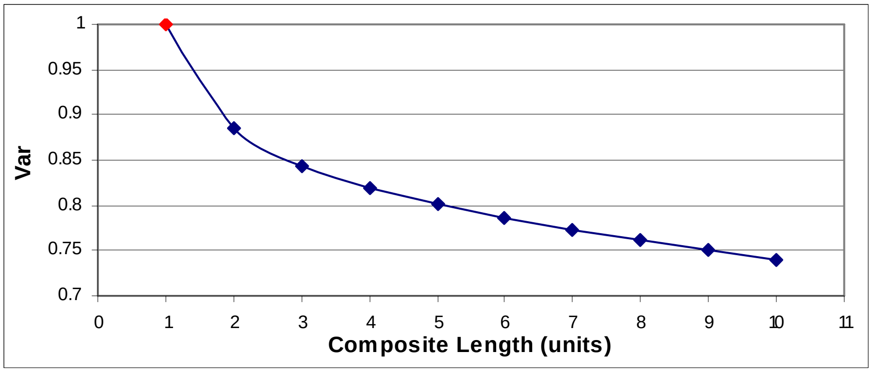

Figure 5: Variance of composites at different lengths.

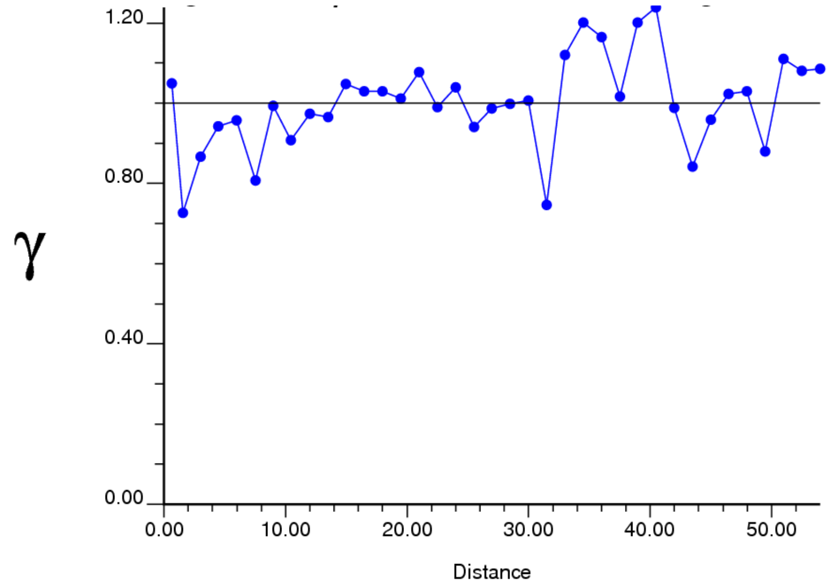

Figure 6: Experimental down-the-hole variogram calculated with gold assays from a disseminated deposit.

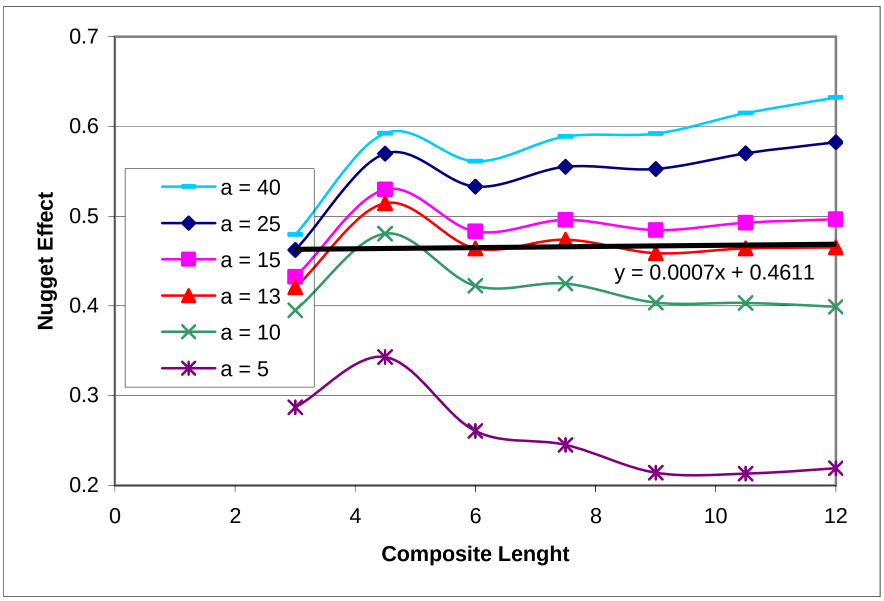

Figure 7: Estimated gold nugget effect values at different composite lengths and with different tentative variogram ranges.

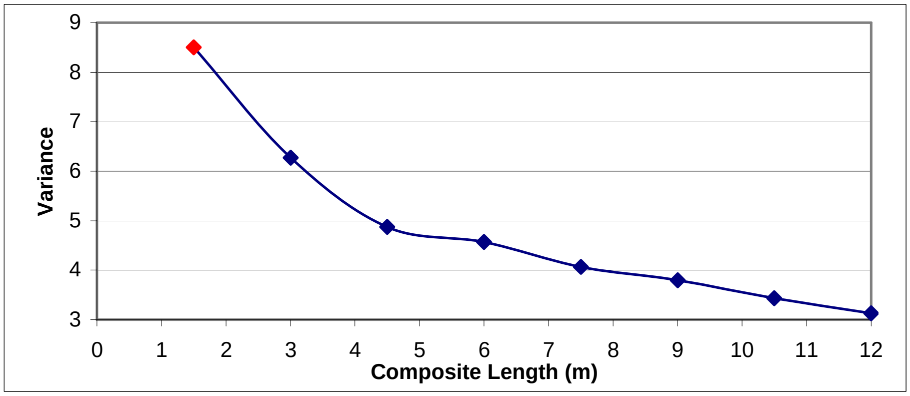

Figure 8: Variance of gold samples composites at different composite lengths.

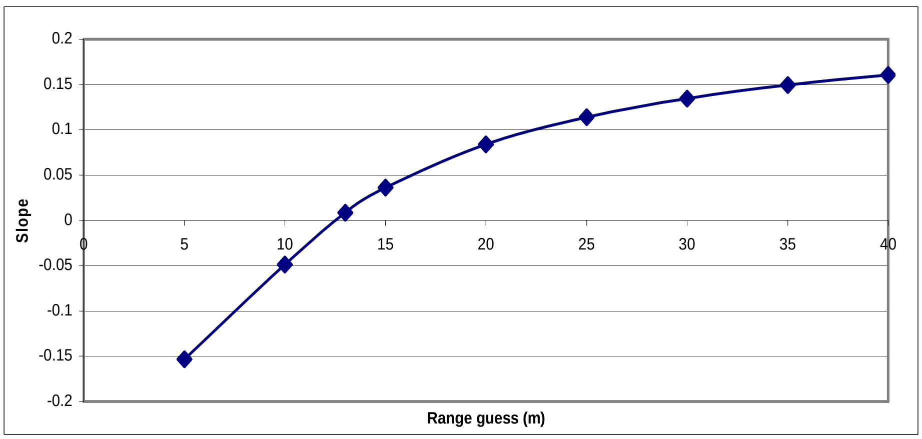

Figure 9: Change in the slope of the inferred nugget effect vs. composite length regression line, at different variogram ranges for gold samples.

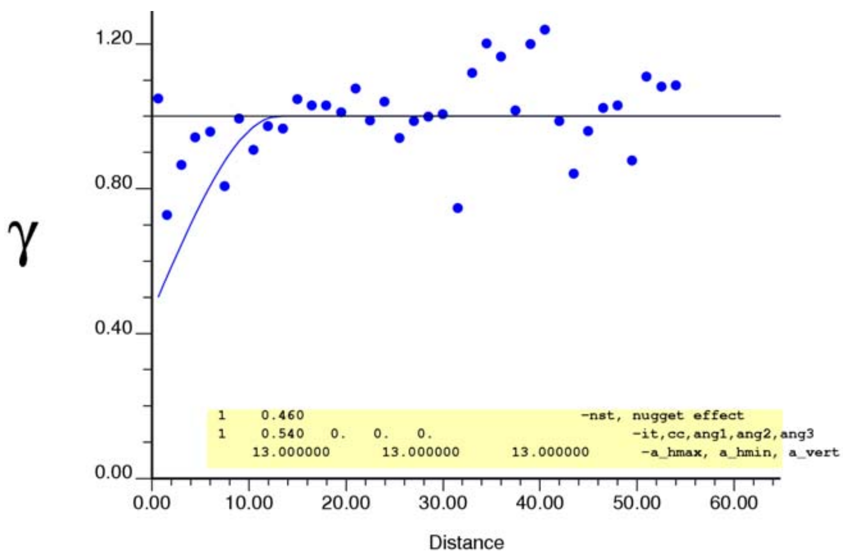

Figure 10: Down-the-hole gold variogram model fitted with the optimized nugget effect and variogram range parameters

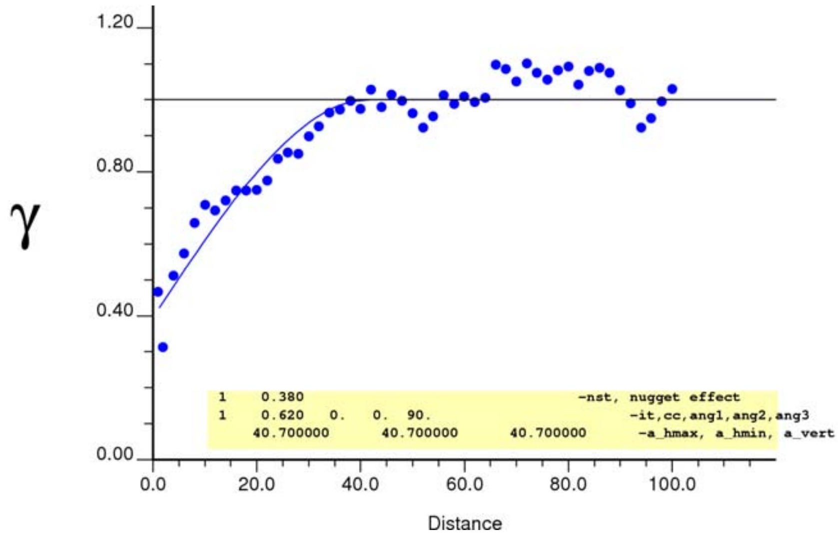

Figure 11: Copper down-the-hole experimental variogram and the correspondent fitted model using the calibrated parameters.

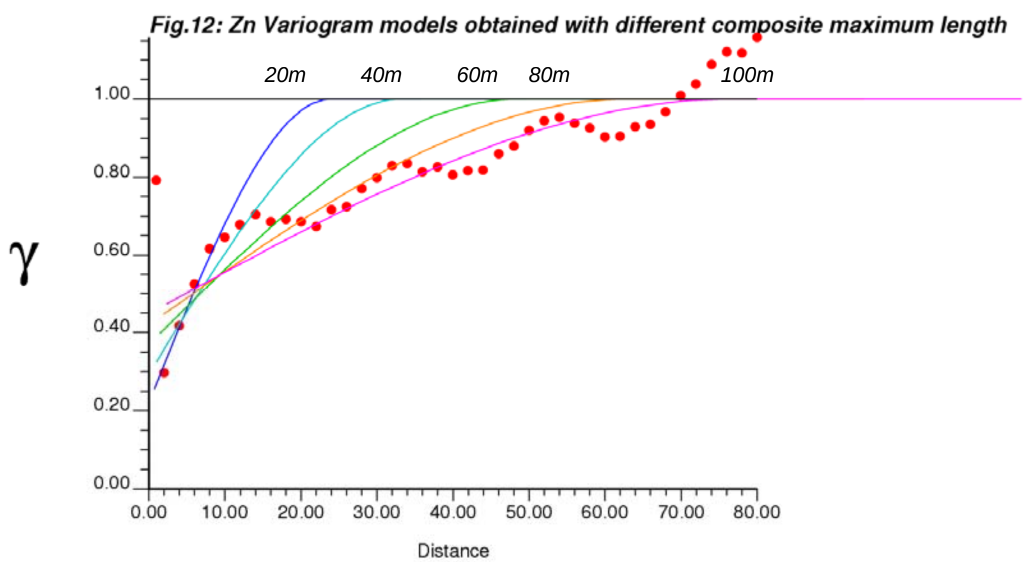

Figure 12: Zn variogram models with parameters obtained using different inputs in the maximum composite length.