Compositing for the Nugget Effect

Mahshid Babakhani and Clayton V. Deutsch

The nugget effect is an important property of the variogram. The nugget effect has a large impact on the outcome of kriging and simulation. The variance of the original sample data and the variance of the data composited to different scales are used in order to calculate the nugget effect at different scales.

Introduction

Though the value of the variogram for $h = 0$ is strictly 0, several factors, such as sampling error and short scale variability cause the variogram to increase abruptly. This causes a discontinuity at the origin of the variogram called the nugget effect (Isaaks, 1989). Understanding the nugget effect can provide important information on the phenomenon under study (Carrasco, 2010). The nugget effect represents the randomness of the variable at $h \rightarrow 0$, which is a characteristic of the variable. It is isotropic or constant in all directions (Pincock Perspectives, 2007). Estimating the nugget effect from the best informed direction (downhole variogram) is a good idea, but it could be unclear and another approach would provide supporting information.

In practice, if the drill-holes have equal-length data are available; the nugget effect can preferably be calculated by using the variance of the sample data and composited data. The idea of this paper is to find the nugget effect by calculating the variance of the composited drill-hole data for different composite scale and the variance of the original data based on the number of data that were included in the compositing progress. Then it will be shown that for reasonable range the practical nugget effect for appropriate composite number is close to the true point scale nugget effect.

Implementation

There is a relationship between the nugget effect of the sample data and composited data. If the original data are at the scale of *v*, consider the averaging to larger scale *V* and without loss of generality assume that *V=nv*. If the sample data are Z_i , so the composited data are defined as below:

$$Z_V = \frac{1}{n}\sum_{i=1}^{n} Z_i \tag{1}$$

Also it is assumed that the data are random as pure nugget effect and stationary so the variance is equal to nugget effect. With this assumption the covariances between the samples within the composites can be written as:

$$\begin{aligned} Cov(Z_i, Z_j) &= \sigma_v^2 &&if\ i = j \\ Cov(Z_i, Z_j) &= 0 &&if\ i \neq j \end{aligned} \tag{2}$$

The variance of sample composite can be calculated by:

$$\sigma_V^2 = Var(Z_V) = \frac{\sigma_v^2}{n} \tag{3}$$

Equation (3) shows that the nugget effect drops very quickly when the data are composited, so it would be fine to do the compositing as much as reasonable. The two variograms in Figure 1 are illustrated to show the accuracy of this equation.

The point scale nugget effect can be estimated by compositing. The variance of the original data and the composite data are the main factor in computing the nugget effect. In the standardized variogram, the variance at the point scale can be written as equation (4):

$$\sigma_v^2 = C_0 + C_1 \tag{4}$$

Where C_0 is the point scale nugget effect, C_1 is the variogram differential sill. In the same way the variance at the composite of the larger scale of *V* is shown in equation (5):

$$\sigma_V^2 = \frac{C_0}{n} + C_1 \tag{5}$$

Where based on the equation (3), $\frac{C_0}{n}$ is the nugget effect at the larger scale of *V* with the composite number of n, C_1 is the variogram differential sill. This is assumed the same as the C_1 at the point scale variogram, because the important but reasonable assumption is that the structural (spatially correlated) variance does not change for reasonable composite lengths. The point scale nugget effect can be calculated by adding these two equations:

$$C_0 = (\sigma_v^2 - \sigma_V^2)(\frac{n}{n-1}) \tag{6}$$

To calculate the point scale nugget effect based on Equation (6), the algorithm is implemented in FORTRAN program `nugg_comp`. In this program it is assumed that the data are in constant length. The data are composited for each drillhole individually and are merged at the end of the data file. If the point scale data is at the length of *v* the composite number multiplies it and the composited data are at the scale of *V=nv*. The values in each of these new scale assays are derived from the averaging of the values in the point scale assays. It is possible that in some of the drillholes some additional data do not consider in compositing progress. For example if there is a sample of five assays and the composite number is equal to 2, it is obvious that the fifth assay is not used in compositing progress. These extra samples are not considered in calculating the composite variance. In order to calculate the variance of the point scale, these data are also eliminated from data file so that the number of data will be the same in calculating the both variances. There are two outputs for this program. One of them is the report file that has the information about the number of data before and after elimination of extra data, the variance in original sample without the extra data and the variance of composited data and of course the nugget effect in both scales. The other output is the data file includes the drillholes ID, X, Y, Z coordinates and composite values. The parameter file for this program is shown in Table 1:

Table 1: Parameters file for the `nugg_comp` program

```
          Parameters for nuggcomposite
          ****************************

START OF PARAMETERS:

../data/cluster.dat    -  file with data

1 2 3 4 5              -  columns ids for ddhid,x,y,z,variable

2                     -  composite number

report.out            -  report file for nugget effects and variances

composite.out         -  output file with composites
```

The point scale nugget effect that is derived from this method should remain the same for the different composite number of n in compositing. The copper values with 10596 data and 135 drillholes ID in three directions of X, Y and Z are used to demonstrate that the point nugget effect derived from different composite numbers remains approximately the same. In Figure 2 it can be seen that the differences in the point nugget effect values are slight and it can be connived.

Discussion

The variances derived from data composited are sensitive to the short scale variogram that is why the range of the variogram plays an important role on calculating the point scale nugget effect. It can be shown that for small ranges the practical nugget effect is so different from the true point nugget effect. So it is critical to choose a reasonable range for the variogram modeling. To understand more about the effect of the range on point scale nugget effect, Figure 3 is demonstrated the practical nugget effect of the 1000 standard normal unconditional realizations in comparison with the ratio of variogram range to composite length.

In this practice the range is changed from 1 to 50 and the composite number is chosen to be 2. The original length of data assay is equal to 1, so the composite number is equal to composite length. So every 2 data

from the sample data are averaged and put in one assay in composited dataset. It can be seen that at the range of 40, the practical nugget effect is almost the same as true point nugget effect. The large nugget effect for small range indicates that most of the spatial correlation is confined to distances shorter than the sampling interval (Walley, Fu, Groenugen, & Kessel, 2001).

Example
The mining data `lgp.dat` that consists of 10596 samples and 135 drillholes is used to test the program. The Copper variable is the subject of this practice. The `nugg_comp` program calculates the both output file. The report file in Figure 4 shows that the number of involved data in two separate composites number of 2 and 3 are completely different. As it mentioned in previous part, the point nugget effect is almost the same. As it is expected the mean in each compositing is exactly the same and approximately the same in different compositing number.
In Figure 6 the three variograms indicate the traditional variogram in 2 direction of X and Y for the sample data and compositing data and Figure 7 shows the traditional variogram in the direction of Z that is the direction of drillhole. In the dataset only the Z direction is changing with the copper variable inside the drillholes and the other two directions are only change within the drillholes not inside them.

Conclusion
Using the compositing to derive the point scale nugget effect is a simple way because variance of the sample data and the composited data are the only parameters involved in this calculation. Since the variances are coming from the same number of data, the error in calculating the point nugget effect is small. The program `nugg_comp` is written for data of constant length, without missing values and sorted in order down the drillholes.

References
Carrasco, P. C. (2010). Nugget effect, artificial or natural. *The Journal of the Southern African Institute of Mining and Metallurgy*, 299-305.
Isaaks, E. H. (1989). *An introduction to applied geostatistics.* New York: Oxford University Press.
Pincock Perspectives. (2007, May). Consultants for Mining and Financial Solutions. *Basic of Variogram Analysis* p. 5.
Walley, F., Fu, G., Groenugen, J.-W., & Kessel, C. (2001). Short-Range Spatial Variability of Nitrogen Fixation by Field-Grown Chickpea. *Soil Science Society of America Journal*, 1719.
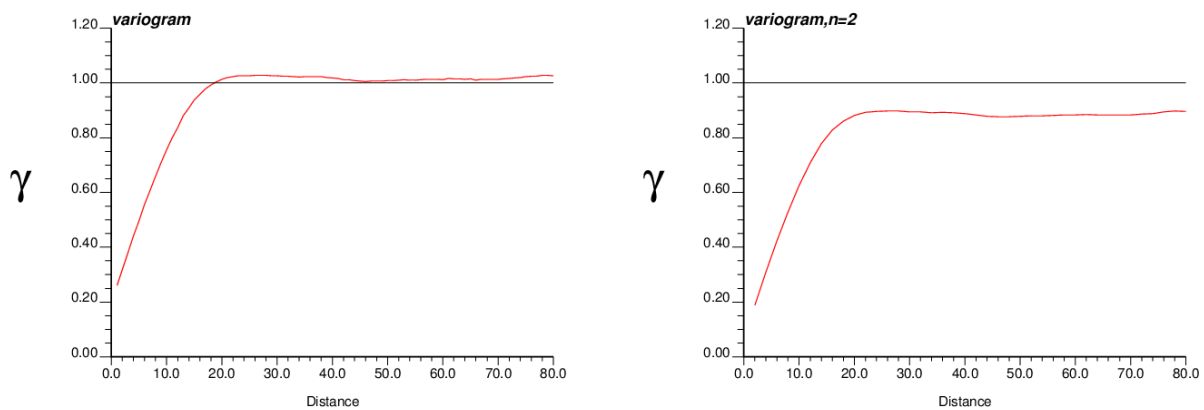
Figure 1: The left variogram is a variogram with nugget effect of 0.25 and the right one is the variogram of the composite data with composite number of two and the nugget effect of 0.12
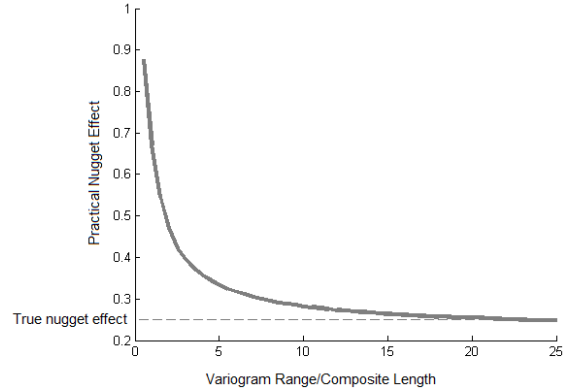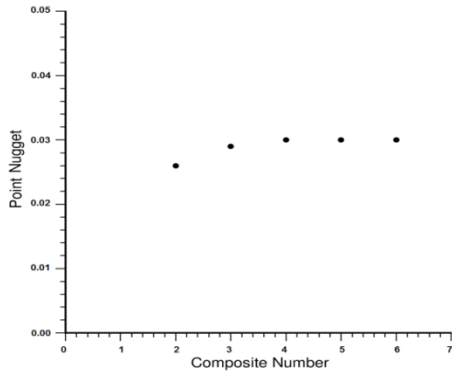
Figure 2 : The point nugget effect for different composite number.   **Figure 3:** The relationship between the variogram range and the nugget effect

```
Composite Nugget Report

-----------------------

Point data:                 10596
Point data after comp:      10534
Point variance:             0.071
Composite variance:         0.058
Point mean:                 0.252
Composite mean:             0.252
Point Nugget:               0.026
Composite Nugget:           0.013
```

```
Composite Nugget Report

-----------------------

Point data:                 10596
Point data after comp:      10464
Point variance:             0.072
Composite variance:         0.053
Point mean:                 0.253
Composite mean:             0.253
point Nugget:               0.029
composite Nugget:           0.010
```

Figure 4: the left repot file is for composite number of 2 and the right one is for composite number of 3
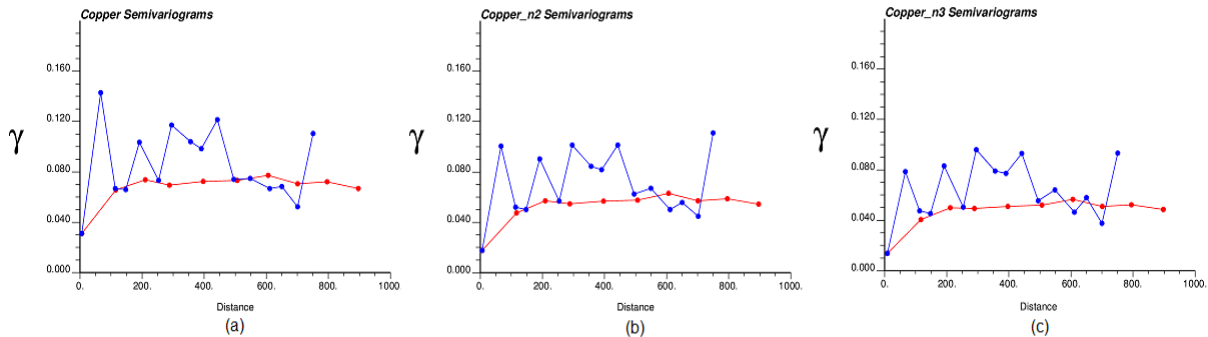

Figure 5: The blue line is in Y direction and the red one is in X direction. Figure (a) is the variogram of the original data, figure(b) is the variogram of composited data with the composite number of 2 and figure (c) is the variogram of composited data with the composite number of 3
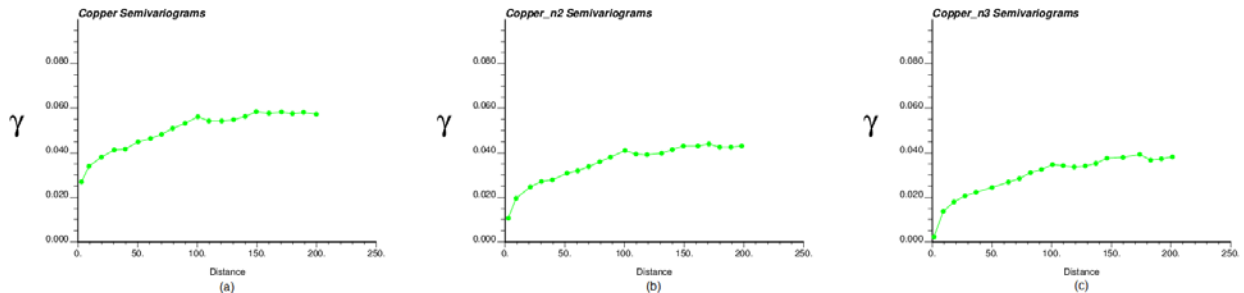

Figure 6: Figure (a) is the variogram of the original data, figure(b) is the variogram of composited data with the composite number of 2 and figure (c) is the variogram of composited data with the composite number of 3