A test of two methods of simulation of regionalized variables

by HARI S. PANDALAI and MARGARET ARMSTRONG, Associate Member

ABSTRACT

Two methods of Monte-Carlo simulations, one suggested by Matheron (1972) and the other by Rudenko (1983), have been tested for simulating coal deposits. To do this 2160 values of ash percentages were simulated using 60 known data points in a two dimensional grid, 20 m x 40 m. It is found that while both methods conserve the mean of the known distribution in the results of the simulations, only the first method reproduces the spatial covariance (i.e. the variogram), of the experimental data. The results show that for simulation of regionalized variables only the first method tested yields the desired results.

Keywords: Coal, coal ash, conditional simulation, geostatistics, Monte-Carlo simulation.

INTRODUCTION

The aim of this work has been to make a comparative assessment of two methods of Monte-Carlo simulation and to demonstrate practically the efficacy of each with the help of a study on borehole data from the coal seams of the Bowen Basin of Australia. The first method is based on the theory of conditional simulation developed by Matheron (1972), while the second is based on a method suggested by Rudenko (1983). Both methods were designed to provide numerical models of deposits which can then be "exploited" using different mine plans. So they must be able to duplicate the main features of the deposit.

The method of conditional simulations developed by Matheron (1972), Journel (1974), and Journel and Huijbregts (1978), aims at generating a set of values with the same geostatistical characteristics as that of the known data points i.e. the same mean, variance and variogram. Moreover the simulation has to pass through the experimental points. The objective here is therefore to obtain a numerical model Zr(x) of the deposit which has the same geostatistical characteristics as Z(x), the data, which is itself considered as a realisation of the regionalised variable Z(x). This ensures that the numerical model Zr(x) is also a realisation of the same, regionalised variable Z(x).

The method essentially consists of summing the value kriged at the point to be simulated with an "error of kriging" which is isomorphous to the actual error of estimation at the point. The simulation uses the relationship

\[ Zr(x) = Zk(x) + (Zr(x) - Zk(x)) \]

where \( Zr(x) \) is the conditionally simulated value at the point \( x \) \( Zk(x) \) is the value kriged at point \( x \) using known data points \( Zr(x) \) is the non-conditionally simulated value at the point \( x \) and \( Zr(x) \) is the value kriged at point \( x \) using values of \( Zr(x) \) at the known data points.

The term \( (Zr(x) - Zr(x)) \) is considered isomorphous to the actual error of estimation at the point \( x \) on the hypothesis that the non-conditional simulations \( Zr(x) \) are the realisations of the regionalised variable \( Z(x) \). The method relies basically on the generation of a set of values \( Zr(x) \) in such a manner that they possess the same spatial covariance as that of the known data values. This is achieved by the method of turning bands (Matheron, 1972).

The second simulation method suggested by Rudenko (1983) is simpler and involves summing the kriged values at a point (or the kriged mean value of a block) with the standard deviation of estimation (of a point or a block as the case may be) multiplied by a random number with a \( \tau \) distribution. Thus for simulation of values at points (for example),

\[ Zr(x) = Zk(x) + (S(x) \times \eta) \]

where \( Zr(x) \) is the value simulated at point \( x \), \( Zk(x) \) is the value kriged at point \( x \) using known data values, \( S(x) \) is the square root of the kriging variance at point \( x \), \( \eta \) is a random number with a \( \tau \) distribution with one fewer degrees of freedom than the number of parameters samples.

Several repetitions of simulations at the points are calculated and the mean is taken as the final simulated value. Rudenko (1983) examined the simulations using (1) a different random number for each point and several repetitions to obtain the final value and (2) one random number for the entire set of points for one simulation followed by several such simulations to obtain the final value. He concluded that the second method was more suitable.

METHODOLOGY OF CONDITIONAL SIMULATION

This method can be applied directly to the known (conditioning) data or to a gaussian transform of the data. However, the results of the simulation are usually better when using the transformed data because the method involves the simulation of several large number of random numbers with a gaussian distribution to obtain the non-conditionally simulated values at each point with the predetermined spatial covariance (Journel and Huijbregts, 1978). The transformation of the data is carried out with Hermite polynomials using the relation

\[ Z(x) = \phi(Y(x)) = \sum \phi_n H_n(Y(x)) \]

\[ n = 0 \]

\[ Z(x) = \phi(Y(x)) = \sum \phi_n H_n(Y(x)) \]

\[ n = 1 \]

where variance \( (Z(x)) = \sum (\phi_n) \)
$Z(x)$ is the data value at point $x$
$Y(x)$ is the transformed data, $N(0,1)$
$\phi$ are the coefficients of the Hermite polynomials, and
$H$ are the Hermite polynomials.

Hermite polynomials are used because they form an orthonormal basis relative to the normal (i.e. Gaussian distribution). For more information about these, readers can consult Kendall and Stuart (1960), Abramowitz and Stegun (1964) or Szego (1939).

The variograms are calculated and modelled for the original data points. The variogram of the transformed data is also calculated and modelled using the same model, except that the sills of the structures are adjusted so as to obtain a total variance of 1.0 for the Gaussian transform.

The simulation is done with the transformed values $Y(x)$ using the corresponding variogram and finally the result is back-transformed to obtain the desired simulated values.

The work was done at the Centre de Géostatistique at Morphologie Mathématique.

**PRACTICAL TEST OF THE SIMULATION METHODS**

The present study was carried out on a set of values representing the ash accumulation (sum of ash per cent weighted by the thickness of sample) in a selected area of the coalfield which had been sampled by 60 bore-holes spaced roughly in grid of 150 m × 200 m. The data values had a mean accumulation of 35.31 and a variance equal to 35.13.

For the purpose of the demonstration it was considered sufficient to simulate point values at a grid interval of 20 m × 40 m over a total area of about 940 m × 1800 m (i.e. 48 × 45 points).

The semivariogram ($\gamma(h)$) of $Z(x)$ was calculated and modelled as the sum of two spherical structures with ranges of 250 m and 920 m and with sills of 22.5 and 12.63 respectively. The corresponding semivariogram model for $Y(x)$ was found to be the sum of two spherical structures with sills of 0.68 and 0.32 and the same ranges as before.

For the conditional simulation, each point to be simulated was kriged at all the data points using the variogram of $Y(x)$, firstly with the transformed values ($Y(x)$) and then with the non-conditionally simulated values ($Y(x)$) to obtain $Y_k$ and $Y_{sk}$ (corresponding to $Z_k$ and $Z_{sk}$ respectively in the equation for $Z_{sk}$).

In the simulation method proposed by Rudenko (1983), the variogram of $Z(x)$ was used and each point to be simulated was kriged using values of $Z(x)$ at all the data points. Because of the large number of points simulated it was considered sufficient to use random numbers with a gaussian distribution between −5.0 and +5.0 instead of the ‘t’ distribution suggested by Rudenko.

One hundred repetitive simulations were performed to calculate the final simulated value in methods 1 and 2.

The statistics of the 2160 points simulated are represented in Table 1.

**Table 1**

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Mean</th>
<th>Variance</th>
<th>$S1$</th>
<th>$S2$</th>
<th>$S3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>34.98</td>
<td>35.09</td>
<td>1.0</td>
<td>0.73</td>
<td>0.72</td>
</tr>
<tr>
<td>S2</td>
<td>35.06</td>
<td>17.96</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>S3</td>
<td>35.01</td>
<td>17.91</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

It is clear that although the values of the conditional simulation have approximately the same variance as the experimental data they are supposed to reproduce, those obtained using Rudenko's method are much less variable.

Figure 1 shows the distribution of the original data compared with the results ($S1$) of the conditional simulation and that ($S2$) of method (1) proposed by Rudenko. The distribution of values ($S3$) generated by method (2) was identical to that of $S2$. As expected the kriging used in methods 1 and 2 of Rudenko conserves the mean which is not changed by the addition of the product in the latter part of the equation of simulation whose expectation is theoretically zero. The variance of the data $S2$ and $S3$ generated was not modified much from the variance of the kriged values as the variance of the product $SK(x) \times m$ is small and the covariance between $Zk(x)$ and $(SK(x) \times m)$ is zero. The variograms of the simulated data $S1$ and $S2$ are presented in Fig. 2 together with the modelled variogram of $Z(x)$. These quite
clearly show that the S2 do not have the same spatial distribution as the experimental data Z(x). (The variogram of S3 was found to be identical to that of S2.) Indeed the variance of S2 and S3 could be increased to obtain the variance of the original data by increasing the variance of the distribution of random numbers but then the varigrams of the simulated values data would agree even less with that of Z(x).

The utilisation of the numerical models of deposits obtained by conditional simulation in planning cycles of extraction, stockpiling and treatment of ores, have been extensively dealt with by Deraisme and Dumay (1979, 1981) and Deraisme and de Fouquet (1984). The fluctuations in point grades obtained in an hypothetical cycle of mining with the results S1 and S2 are shown in Fig. 3. The simulation S2 is clearly much less variable than S1. This occurs because the kriging variance is much less than the actual variance of the data. So S2 gives an entirely false impression of the homogeneity of the run of mine coal and could lead mine planners to make serious errors in deciding whether to blend or stockpile the coal. Planning mining operations using the numerical model S1 which has the same geostatistical characteristics as the regionalised variable Z(X) (i.e. which can be considered as a realisation of the regionalised variable Z(X)) would lead to much more realistic conclusions.

It should also be stressed that in the example dealt with in this article only simulations of point values have been done. Simulation of block values entails the 'change of support' (see Matheron, 1963; Journel, 1978) of the variable Z(x) (which is defined on the support of the borehole point) to that of the block of the particular dimensions and the necessary calculations therein.

CONCLUSION

It is evident that for the purpose of simulation where the numerical model must have the same geostatistical characteristics as the experimental data Z(x) only the method of conditional simulation proposed by Matheron (1972) achieves the desired results. Any planning (mining or economic) based on numerical models which do not represent realisations of the regionalised variable Z(X) is likely to lead to erroneous predictions.

REFERENCES

Deraisme, J. and Dumay, R., 1981. Simulations sur modèles géostatis-
tiques du cycle de production et de valorisation des minerais. Inter-
nal report of the Centre de Géostatistique et Morphologie Mathématique, N-680.
Journel, A. G., 1974. Geostatistics for conditional simulation of ore-
bodies, Economic Geology, 79: 673-687.

Fig. 3—(a) Simulated ash accumulations obtained from a condi-
tional simulation.
(b) Simulated ash accumulations obtained using Rudenko's method.