Geostatistical Conditional Fractal Simulation With Irregularly Spaced Data

D.Kentwell, L.Bloom & G.Comber

School of Computer, Information and Mathematical Sciences
Edith Cowan University - Perth Western Australia

Abstract Using the model of fractional Brownian motion a method is given for carrying out conditional simulations on a sparse irregularly spaced data set. The method, on average, maintains not only the fractal co-dimension but also the histogram, mean, variance and spatial correlation of a two dimensional random field. The method can be implemented in the same way as either sequential Gaussian simulation or LU decomposition and does not require the use of spectral functions. A sequential Gaussian example is given using the widely published Berea sandstone data set.

1. INTRODUCTION

Fractal simulations of two dimensional surfaces or random fields are typically carried out with either some form of midpoint displacement method or some form of Fourier transform method. Although both of these methods and their variants maintain the fractal co-dimension of the field they do not always maintain the spatial correlation and are not always conditional. In addition the variants of the midpoint displacement method do not handle irregularly spaced conditioning data. On the other hand geostatistical simulation methods such as sequential Gaussian simulation and LU decomposition, while taking no account of the fractal co-dimension, do maintain a specific spatial correlation, are conditional, and work with irregularly spaced data.

The method proposed here is based on a fractal interpolation algorithm detailed by Rümelin in two papers in 1990 and 1992. His algorithm uses the *covariance of the increments* of fractional Brownian motion together with what are essentially kriging equations to simulate a fractal field that maintains both the fractal co-dimension and its associated spatial correlation. This algorithm provides estimates and error variances that are the same as ordinary kriging with a power model but without having to resort to the use of a pseudo-covariance model and without the use of the Lagrange parameter to solve the kriging system.

2. THE FRACTAL CO-DIMENSION AND FRACTIONAL BROWNIAN MOTION

The increments of fractional Brownian motion $B_H(\mathbf{u}_\alpha) - B_H(\mathbf{u}_\beta)$ (Mandelbrot & Wallis, 1969) in any number of Euclidean dimensions, have a Gaussian distribution with variance

$$E[|B_H(\mathbf{u}_\alpha) - B_H(\mathbf{u}_\beta)|^2] = V_H|\mathbf{u}_\alpha - \mathbf{u}_\beta|^{2H} \quad (1)$$

where V_H is a constant of proportionality and $0 < H < 1$ is the fractal co-dimension. Note that $H = 0.5$ gives the traditional Brownian motion with $\Delta B(\mathbf{u})^2 \propto |\mathbf{h}|$ where $\mathbf{h}$ is any increment vector $\mathbf{u}_\alpha - \mathbf{u}_\beta$. The left hand side of (1) is equivalent to a variogram function as it is a variance of increments, hence

$$2\gamma(\mathbf{h}) = V_H|\mathbf{h}|^{2H}$$

or

$$\gamma(\mathbf{h}) = \tfrac{1}{2}V_H|\mathbf{h}|^{2H} \quad (2)$$

and V_H is therefore the total variance at the reference unit scale $|\mathbf{h}|=1$.

2.1 Determining the fractal co-dimension

Equation (2) is the same as the geostatistical power model. Distributions that conform to (2) are statistically self affine since variations over any scale $r|\boldsymbol{h}|$ are related to the variations over scale $|\mathbf{h}|$

by

$$\gamma(r\mathbf{h}) = r^{2H}\gamma(\mathbf{h}). \qquad (3)$$

This implies that the variance at any scale can be determined by the variance measured at any other scale (Hewett 1986). In practice the experimental semi-variogram can be determined from the available data using

$$\gamma(\mathbf{h}) = \frac{1}{2n}\sum_{\alpha=1}^{n}(z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h}))^2 \qquad (4)$$

where n is the number of pairs $(z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h}))$ at lag $\mathbf{h}$. If the experimental semi-variogram is plotted with log scales on both axes the slope of the fitted line is equal to 2H and the anti-log of the $\gamma(\mathbf{h})$ axis intercept is $\frac{1}{2}V_H$, i.e.

$$2H = \Delta \ln \gamma(\mathbf{h})/\Delta \ln |\mathbf{h}| \qquad (5)$$

and

$$\tfrac{1}{2}V_H = e^{\gamma(0)}. \qquad (6)$$

Linear regression is used to find the slope of log-log experimental semi-variograms and thus to determine $2H$. A power model with the appropriate power $2H$ can then be fitted to the experimental semi-variograms in order to determine $\frac{1}{2}V_H$. Note that it is possible to fit a power model to a semi-variogram without using the log-log axes but there are many possible combinations of V_H and $2H$ that appear to fit and simply doing this may not provide the correct fractal co-dimension H.

3. CONDITIONAL FRACTAL SIMULATION

The covariance of increments method proposed by Rümelin (1990) allows generation of an arbitrary number of values of fractional Brownian motion in one step, and follows similar logic to the (non-fractal) LU decomposition simulation method described by Davis (1987) and Alabert (1987) but also isolates one of the conditioning values to use as an arbitrary fixed reference value. This reference value is then used to calculate increments and as a benchmark against which to gauge the scaling parameter required to implement the use of the fractal co-dimension H. A summary of the covariance of increments method is given below, all values are normal score values. Consider the following scheme in matrix notation

$$[z_c^{(I)}(\mathbf{u}_i)] = \lambda[\, z(\mathbf{u}_\alpha)] + \mathbf{S}\mathbf{w},$$

$$\alpha = 1,\ldots,n \quad i = 1,\ldots,N \qquad (7)$$

where $\mathbf{w}$ is a vector of independent random variables with standard normal distribution N(0,1), λ is an N x n matrix of weights that maintains the spatial structure of the field while estimating $z(\mathbf{u}_i)$ and $\mathbf{S}$ is an N x N matrix controlling the rate of random variation necessary for each simulated value. If we now arbitrarily choose a specific conditioning value $z(\mathbf{u}_\eta)$ and reformulate (7) in terms of increments with respect to $\mathbf{u}_\eta$ we obtain

$$[z_c^{(I)}(\mathbf{u}_i) - z(\mathbf{u}_\eta)] = \lambda'[\, z(\mathbf{u}_\alpha) - z(\mathbf{u}_\eta)] + \mathbf{S}\mathbf{w}$$

$$\alpha = 1,\ldots,n-1 \quad i = 1,\ldots N \qquad (8)$$

where λ' is an N x n-1 matrix (λ without the $\mathbf{u}_\eta$ terms). Multiplying both sides of (8) from the right by $[z(\mathbf{u}_\beta) - z(\mathbf{u}_\eta)]^T$ and taking expectations reduces this to

$$[C(\mathbf{u}_i - \mathbf{u}_\alpha)] = \lambda'\,[C(\mathbf{u}_\alpha - \mathbf{u}_\beta)] \qquad (9)$$

This linear system is similar to a kriging system and can be solved for λ' by Cholesky decomposition if the two covariance matrices are known. Having solved for λ', λ can then be found by the inclusion of

$$\lambda_{i\eta} = 1 - \sum_{j=1}^{n-1}\lambda_{ij} \quad i = 1,\ldots,N. \qquad (10)$$

To obtain matrix $\mathbf{S}$ we multiply both sides of equation (8) from the right by $\mathbf{w}^T$ and take expectations giving

$$E[(z_c^{(I)}(\mathbf{u}_i) - z(\mathbf{u}_\eta))\mathbf{w}^T] = \mathbf{S}. \qquad (11)$$

We now multiply both sides of (8) by $[z_c^{(I)}(\mathbf{u}_j) - z(\mathbf{u}_\eta)]^T$ and take expectations and we arrive at

$$\mathbf{S}\,\mathbf{S}^T = [C(\mathbf{u}_i - \mathbf{u}_j)] - \lambda'\,[C(\mathbf{u}_i - \mathbf{u}_\alpha)]^T. \qquad (12)$$

Now $\mathbf{S}$ can be computed by Cholesky decomposition. Rümelin (1990) has shown that

$$[C(\mathbf{u}_\alpha - \mathbf{u}_\beta)] = \tfrac{1}{2}V_H[|\mathbf{u}_\alpha - \mathbf{u}_\eta|^{2H} - |\mathbf{u}_\alpha - \mathbf{u}_\beta|^{2H}$$

$$+|\mathbf{u}_\eta - \mathbf{u}_\beta|^{2H}] \qquad (13)$$

and this applies for any covariance matrix of the increments between any paired combination of $\mathbf{u}_\alpha, \mathbf{u}_\beta, \mathbf{u}_i$ and $\mathbf{u}_j$. (Note, Rümelin (1990) did not include the proportionality constant V_H as we have here). We now have all that is required to calculate the simulation values as in (7). The notation for this computational scheme can be reduced to what is essentially a set of ordinary kriging and error

variance equations without the Lagrange parameter where λ' is the matrix of weights without the weight of the closest point $\mathbf{u}_\eta$, i.e.

$$\mathbf{C}_{\alpha,\beta|\eta}\,\lambda' = \mathbf{C}_{\alpha,i|\eta} \qquad (14)$$

$$SS^T = \mathbf{C}_{i,i|\eta} - \lambda'\,\mathbf{C}^T_{\alpha,i|\eta} \qquad (15)$$

Equation (15) reduces to

$$s^2 = C_{i,i|\eta} - \lambda'\,\mathbf{C}^T_{\alpha,i|\eta} \qquad (16)$$

when only one point at a time is being simulated. The weight for point $\mathbf{u}_\eta$ is then found from (10). From here the point estimate and variance factor can be computed.

This method still has the current restriction common to all LU decomposition methods, that the number of points that can be simulated at any one time is limited to several hundred because of the size of the covariance matrices. If applied on a regular grid with conditioning data also on a regular grid this method is very fast for simulating large numbers of nodes. It can with some care be applied to conditioning data that is not on a regular grid but since a different location geometry occurs for every new simulation location a different linear system must be solved each time. Computationally this is very time consuming if we are following an LU approach simulating many points at one time.

4. SGFRACT

It is proposed here that a sequential approach utilising an adaptation of the covariance of increments method, simulating only one value at a time, can be used to compute a geostatistical simulation provided there are a sufficient number of local irregularly spaced conditioning data. Computationally solving a linear system involving only one simulation point with up to 30 conditioning values is still fast and it is reasonable to recalculate the linear system at every individual simulation point. We will call this new simulation scheme *sequential Gaussian fractal simulation* and call the computer program SGFRACT. The implementation of the scheme parallels that of sequential Gaussian simulation (Deutsch & Journel 1992) except that the parameters of the conditional distribution at each point to be simulated are calculated using the covariance of increments in (13) to form the linear system instead of simple kriging estimate and error variance. The GSLIB (Deutsch & Journel 1992) FORTRAN routine SGSIM was adapted to create SGFRACT and then used to compute the following fractal simulation example. Points to note about the adaptation are:

(a) A feedback loop is included so that nodes not simulated due to lack of close data can be re-examined after the first random path is completed.
(b) An option is included to set the minimum number of data and/or simulated nodes that together are used to simulate any point thus allowing the option for nodes to be simulated entirely from previously simulated nodes without any original conditioning data.
(c) A normalisation factor is applied to the square root of the fractal variance to bring the simulated population variance back to one. (This factor is not fixed and is dependent on the geometry of the particular simulation. For example the normalisation factor changes with the model, the grid size and/or the overall field size.) This normalisation factor is applied at each individual point simulation rather than at the completion of the simulation so that it does not displace the conditioning data values.
(d) Following Rümelin (1990) a Cholesky forward/backward substitution routine is used to solve the linear system as opposed to Gaussian elimination routine as used in SGSIM.
(e) The search radius is explicitly set by the model to equal the range and cannot be altered. This is because when trying to reproduce a sill the power model is only valid up a semi-variance of one. Beyond that no model is used and the spatial structure is uncontrolled. (see section 4.1)

4.1 Fractal Modelling

When using normal score data the concept of range as applied to a power law model is equivalent to the distance at which the model semi-variance equals 1. Therefore the first step when modelling with a power law is to decide on the approximate lag at which any range or flattening occurs. The power $2H$ should then be determined to one decimal place from the slope of the linear regression of the log-log scaled experimental semi-variogram up to where the value of $|\mathbf{h}|$ equals the range (see figure 3). Using this value for $2H$ the power law model is fitted by eye to the experimental semi-variogram by adjusting the coefficient $\frac{1}{2}V_H$. The exact range can then be calculated by

$$a = e^{-\ln(\frac{1}{2}V_H)/2H} \qquad (17)$$

and with normal score data should be calculated to two decimal places. This range then serves as the search radius calculated by SGFRACT. The coefficient $\frac{1}{2}V_H$ could be taken directly from the regression but it is best to check the fit by eye as there may be other factors to be considered, such as model fits in other directions or extremely erratic experimental semi-variogram data. If the semi-variogram being fitted has no sill then the regression and power model should be fitted up to the largest reliable lag, usually half the field size in any given direction.

5. EXAMPLE USING THE BEREA DATA

The *Berea* data (see figure 1) is a real two dimensional data set consisting of 1600 points derived from air permeability measurements taken on a slab of Berea sandstone (Giordano et. al. 1985). Chu & Journel (1994) used this data set to demonstrate a spectral fractal simulation method. They used a random sample of 64 points from the *Berea* data and we do the same. However we do not know the random selection details which they used and our randomly selected points will not necessarily be the same as theirs (see figure 2).

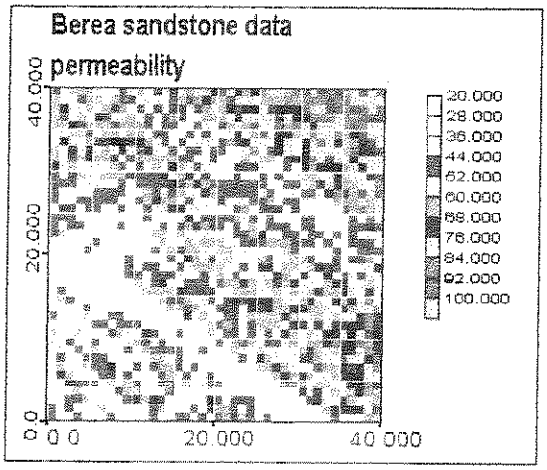


Figure 1: Berea data full set.

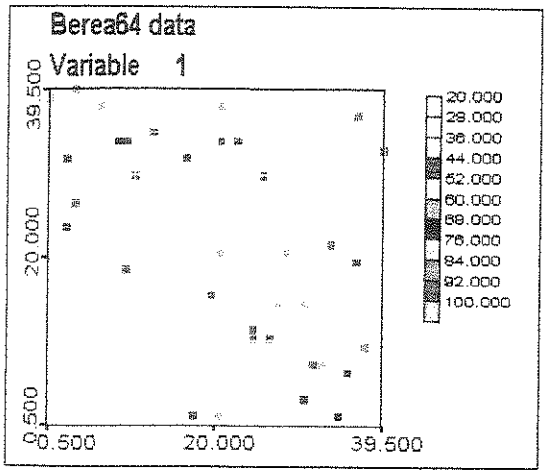


Figure 2: Berea sample data.

For this example we use the variography and normal scores that are derived from the actual *Berea* data rather than the sample data in order to minimise uncertainties while demonstrating SGFRACT.

The anisotropic semi-variogram modelling with fractal power models is shown in figure 3. Note that use of a nugget effect is not necessary. The principal direction of anisotropy is at 123^o where 0^o is north and corresponds to the y axis with angles measured clockwise from 0^o. The model in the 123^o direction is

$$\gamma(\mathbf{h}) = 0.20|\mathbf{h}|^{0.4}$$

which gives a range of 55.9 units. The model in the 33^o direction is

$$\gamma(\mathbf{h}) = 0.38|\mathbf{h}|^{0.4}$$

and gives a range of 11.2 units. We can calculate the anisotropy ratio τ in two ways, by simply taking the ratio of directional ranges given by (17) as is usual or by defining

$$\tau = \nu^{1/2H} \qquad (18)$$

where $\nu = (\frac{1}{2}V_{H1})/(\frac{1}{2}V_{H2})$ and $\frac{1}{2}V_{H1}$ is the coefficient in the principal direction of anisotropy (Chu & Journel 1992). The anisotropy ratio is then 0.20. Figure 4 shows three realisations from the Berea sample data computed by SGFRACT using the anisotropic power model above.

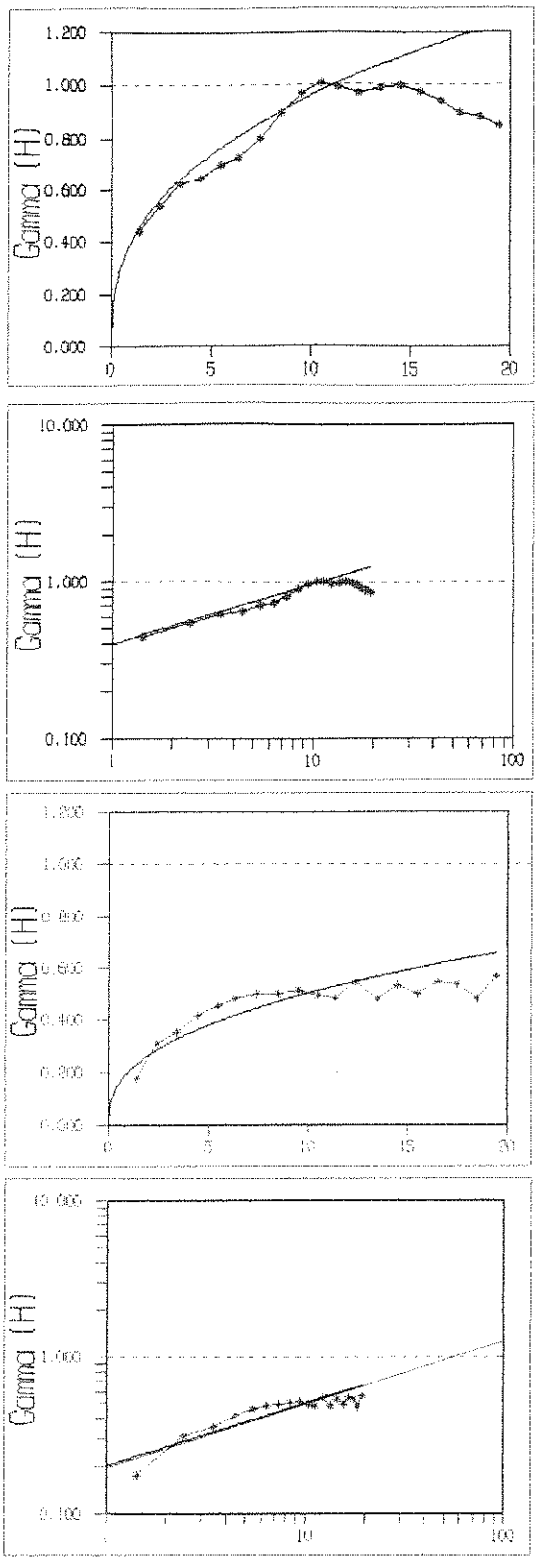
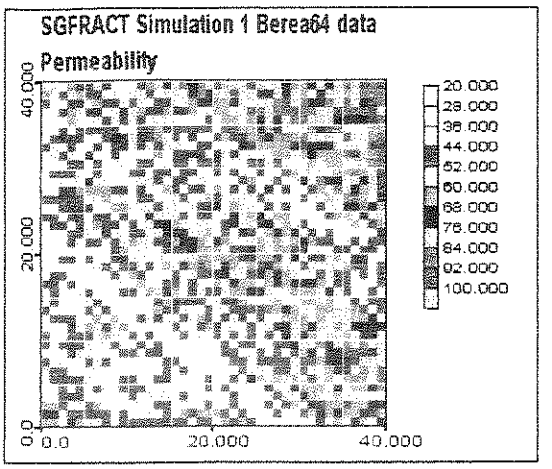
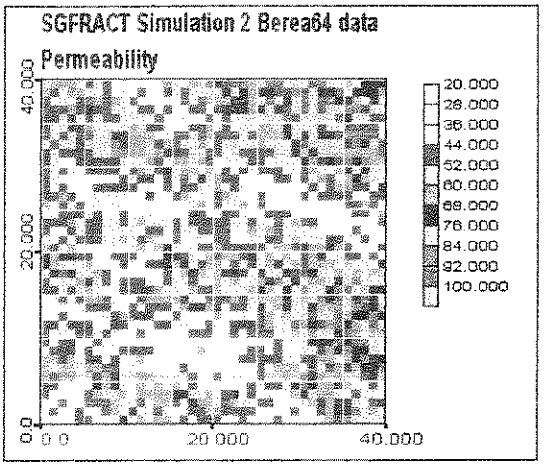


Figure 3: Directional normal score semi-variograms and fitted power models. 33^o (top), 33^o log-log axes (2nd. top), 123^o (2nd. bottom) and 123^o log-log axes (bottom)





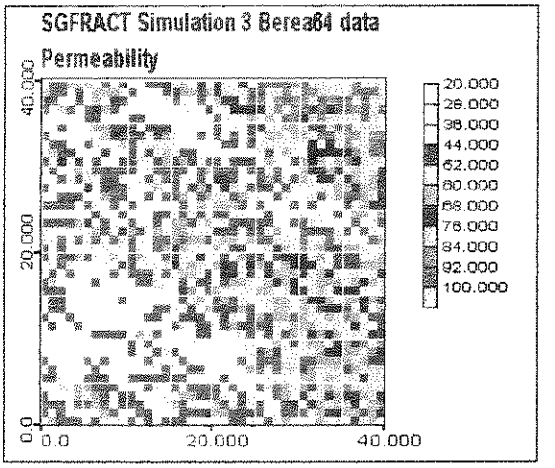


Figure 4. Three selected simulations from the Berea sample data.

6. CONCLUSIONS

The significant contributions of this paper are:
(a) A method for the use of the power model with stationary data.
(b) The ability to carry out conditional fractal simulations using sparse irregularly spaced data without the necessity of spectral functions.
(c) The incorporation of the fractal co-dimension as another statistic that can be reproduced with Gaussian geostatistical simulations.

More work needs to be done on determination of the normalisation factor in the hope that it can be determined theoretically rather than experimentally.

The incorporation of the fractal co-dimension into sequential Gaussian simulations seems to produce improved predictions of grade tonnage curves for gold mineralisation type data sets (Kentwell 1997 & Kentwell et. al. 1997) .

REFERENCES

Alabert, F., The practice of fast conditional simulations through the LU decomposition of the covariance matrix. *Mathematical Geology*. **19**, 5, 369 - 386. 1987.

Chu, J., and A.G. Journel, Conditional fBm simulation with dual kriging. in Dimitrakopoulos, R., (ed.) *Geostatistics For The Next Century*. 407 - 421. Kluwer Academic Publishers. 1994.

Davis, M.W., Production of conditional simulations via the LU decomposition of the covariance matrix. *Mathematical Geology*. **19**, 2, 91 - 98. 1987.

Deutsch, C.V., and A.G. Journel, *GSLIB: Geostatistical Software Library Users Guide*. First Edition. Oxford University Press. 1992.

Giordano, R., S. Slater and K. Mohanty, The effects of permeability variations on flow in porous media. *SPE Paper 14365, 60th SPE Annual Conference*, Las Vegas. 1985.

Hewett, T.A., Fractal distributions of reservoir heterogeneity and their influence on fluid transport. *Paper SPE 15386 presented at the 61st Annual Technical Conference and Exhibition of The Society of Petroleum Engineers*, New Orleans. 1986.

Kentwell, D., *Fractal Relationships and Spatial Distributions in Ore Body Modelling*. Masters Thesis. Edith Cowan University. Perth. 1997.

Kentwell, D., L. Bloom, and G. Comber, Improvements in grade tonnage curve prediction via the inclusion of the fractal co-dimension. Preprint. 1997.

Mandelbrot, B.B., J.R. Wallis, Computer experiments with fractional Gaussian noises. Part 3, Mathematical appendix. *Water Resources Research*. 5, 1, 260 - 267. 1969.

Rümelin, W., Simulation of fractional Brownian Motion. in Peitgen et. al. (eds) *Fractals in the Fundamental and Applied Sciences*. Elsevier. 1990.

Rümelin, W., Fractal interpolation of random fields of fractal Brownian motion. in Encarnacao et. al. (eds) *Fractal Geometry and Computer Graphics*. 122 - 131. Springer Verlag. 1992.

NOTATION AND UNITS

$\mathbf{u}$	Location vector
$z(\mathbf{u}_\alpha)$	Regionalised value
$z(\mathbf{u})$	Regionalised variable
$\mathbf{h}$	Translation vector or increment
$\gamma(\mathbf{h})$	Semi-variogram function
$\lvert\mathbf{h}\rvert$	Increment distance or lag
λ	Weight
τ	Anisotropy ratio
$\mathbf{C}$	Variance covariance matrix
λ	Matrix of weights
$z_c^{(l)}(\mathbf{u})$	Conditional simulation
$\mathbf{w}$	Vector of standard normal score
$\mathbf{r}$	Scaling vector where r_i is not necessarily equal to r_j
$B(\mathbf{u})$	Brownian motion
$B_H(\mathbf{u})$	Fractional Brownian motion
H	Fractal co-dimension
V_H	Proportionality constant
$\mathbf{S}$	Estimation error standard deviation matrix
$\mathbf{u}_\eta$	Increment reference point location
s	Standard deviation
$C(\ )$	Covariance