Exact Downscaling in Geostatistical Applications

Weishan Ren

Center for Computational Geostatistics
Department Civil & Environmental Engineering
University of Alberta

Exact downscaling techniques are demonstrated with three different scales: the point, block, and model scales. The model scale can be any scale between the point and the block scale. The model scale is modeled exactly reproducing the large scale. Proofs of the exact reproduction of block data are provided. Some implementation details are also given.

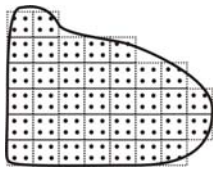
Introduction

One of major tasks of petroleum geostatistics is to build reservoir models of petrophysical properties for reservoir flow simulation. When the target for flow simulation is a small area, downscaling of geostatistical models may be required. In the McMurray Formation where the SAGD technique is commonly used for oil sands recovery, the flow simulation of horizontal well pairs or individual SAGD pads is very important. The detailed flow simulation of selected small areas requires input models at a very fine grid size.

Reliable geostatistical models can be constructed at a large scale over lease areas. Using the large scale model to generate a fine scale model is a downscaling process. Co-kriging, collocated co-kriging, and trend modeling techniques can be used to generate fine scale models; however, the fine scale models are not perfectly consistent with the large scale model. Exact downscaling techniques are introduced to generate fine scale heterogeneous models. These techniques are developed using the direct kriging and Direct Sequential Simulation (DSS) techniques.

The background and theory of the downscaling methods are presented. Some small examples are provided to illustrate the theoretical and practical aspects of the techniques. Detailed implementation of the downscaling methods will also be demonstrated.

Background



Consider a block volume V with an arbitrary shape as shown to the right. The block is composed of a great number of points. If a property is scaled lineally, the block value Z_V can be calculated from the point data by this equation:

$$Z_V = \frac{1}{V}\int Z_\bullet(\mathbf{u})d\,\mathbf{u} \qquad (1)$$

where $Z_\bullet(\mathbf{u})$ is the point datum at any location $\mathbf{u}$ in the block.

If the block can be discretized into *n* equal-sized cells with small volume *v*, the block value can be expressed as a simple linear equation:

$$Z_V = \frac{1}{n}\sum_{i=1}^{n} Z_v(\mathbf{u}_i) \tag{2}$$

where $Z_v(\mathbf{u}_i)$ (i = 1…n) are the values in the small cells. If the small volume is composed of m points, the cell value $Z_v(\mathbf{u}_i)$ can be expressed as:

$$Z_v(\mathbf{u}_i) = \frac{1}{m}\sum_{j=1}^{m} Z_\bullet(\mathbf{u}'_j) \tag{3}$$

where $Z_\bullet(\mathbf{u}'_j)$ (j = 1…m) are the point data in the small cell.

Now consider a large domain A (or a large model area) in which the multiscale data are stationary. Thus, the statistical properties such as mean, variance, covariance of data at each scale are the same over the entire area; and the statistical properties of data at different scales follow same relationships over the entire area. If the scale relationship is linear, the mean is same for all the scales. The data variance decreases as the scale increases. The volume averaged covariances apply to any scales.

In the large model area, the block data are exhaustively available for the whole area. A certain amount of point data is also available. Statistical properties at the block and point scales can be calculated from the available data. We are also interested in the statistical properties at the model scale, which can be any scale between the block and point scales. The mean at the model scale can be inferred from the block data. Because the block data covers the whole model area, the mean of block data can be used for all the scales:

$$\overline{Z}_V = \overline{Z}_v = \overline{Z}_\bullet \tag{4}$$

where the subscript V represents the block scale, v represents the model scale, and • represents the point scale.

Any variance is a dispersion variance of the data in a large volume. The variance of data at the model scale is the dispersion variance of the data in the large domain A, $\sigma_v^2 = D^2(v, A)$. Similarly, the variance of point data is: $\sigma_\bullet^2 = D^2(\bullet, A)$. Krige's Relation gives the variance at the model scale:

$$D^2(v, A) = D^2(\bullet, A) - D^2(\bullet, v) \tag{5-5}$$

where the dispersion variance of the point data in the model scale volume can be calculated from the volume-variance relation:

$$D^2(\bullet, v) = C_{\bullet\bullet}(0) - \overline{C}_{vv}(0) = \sigma_\bullet^2 - \overline{C}_{vv}(0) \tag{6}$$

where the $C_{\bullet\bullet}(0)$ and $\overline{C}_{vv}(0)$ are the covariances of datum itself at the point and model scales. Therefore, the variance at the model scale is actually the volume-averaged covariance at the model scale:

$$\sigma_v^2 = \overline{C}_{vv} \tag{7}$$

The covariance at the model scale can be inferred from the covariance at point data through the volume-averaged covariance relationship.

$$
\begin{aligned}
\overline{C}_{vv}(\mathbf{h}) &= E\{Z_v(\mathbf{u}) \times Z_v(\mathbf{u}+\mathbf{h})\} - E\{Z_v(\mathbf{u})\} E\{Z_v(\mathbf{u}+\mathbf{h})\} \\
&= E\left\{ \frac{1}{|v|} \int_{v(\mathbf{u})} Z_\bullet(\mathbf{x}) d\,\mathbf{x} \times \frac{1}{|v|} \int_{v(\mathbf{u}+\mathbf{h})} Z_\bullet(\mathbf{y}) d\,\mathbf{y} \right\} - m^2 \\
&= E\left\{ \frac{1}{|v||v|} \int_{v(\mathbf{u})} \int_{v(\mathbf{u}+\mathbf{h})} Z_\bullet(\mathbf{x}) Z_\bullet(\mathbf{y}) d\,\mathbf{x}\, d\,\mathbf{y} \right\} - m^2 \\
&= \frac{1}{|v||v|} \int_{v(\mathbf{u})} \int_{v(\mathbf{u}+\mathbf{h})} (E\{Z_\bullet(\mathbf{x}) Z_\bullet(\mathbf{y})\} - m^2) d\,\mathbf{x}\, d\,\mathbf{y} \\
&= \frac{1}{|v|^2} \int_{v(\mathbf{u})} \int_{v(\mathbf{u}+\mathbf{h})} C_{\bullet\bullet}(\mathbf{x}-\mathbf{y}) d\,\mathbf{x}\, d\,\mathbf{y} \\
&= \frac{1}{|v|^2} \sum_{v(\mathbf{u})} \sum_{v(\mathbf{u}+\mathbf{h})} C_{\bullet\bullet}(\mathbf{h}')
\end{aligned}
\tag{5-8}
$$

Similarly, we can get the block scale covariance, block-model covariance, block-point covariance, and model-point covariance so that a total of five covariance models from the point scale covariance.

The block scale covariance:

$$
\overline{C}_{VV}(\mathbf{h}) = E\{Z_V(\mathbf{u}) \times Z_V(\mathbf{u}+\mathbf{h})\} - E\{Z_V(\mathbf{u})\} E\{Z_V(\mathbf{u}+\mathbf{h})\} = \frac{1}{|V|^2} \sum_{V(\mathbf{u})} \sum_{V(\mathbf{u}+\mathbf{h})} C_{\bullet\bullet}(\mathbf{h}') \tag{9}
$$

The block-model covariance:

$$
\overline{C}_{Vv}(\mathbf{h}) = E\{Z_V(\mathbf{u}) \cdot Z_v(\mathbf{u}+\mathbf{h})\} - E\{Z_V(\mathbf{u})\} E\{Z_v(\mathbf{u}+\mathbf{h})\} = \frac{1}{|V||v|} \sum_{V(\mathbf{u})} \sum_{v(\mathbf{u}+\mathbf{h})} C_{\bullet\bullet}(\mathbf{h}') \tag{10}
$$

The block-point covariance:

$$
\overline{C}_{V\bullet}(\mathbf{h}) = E\{Z_V(\mathbf{u}) \cdot Z_\bullet(\mathbf{u}+\mathbf{h})\} - E\{Z_V(\mathbf{u})\} E\{Z_\bullet(\mathbf{u}+\mathbf{h})\} = \frac{1}{|V|} \sum_{V(\mathbf{u})} C_{\bullet\bullet}(\mathbf{h}') \tag{11}
$$

And the model-point covariance:

$$
C_{v\bullet}(\mathbf{h}) = E\{Z_v(\mathbf{u}) \times Z_\bullet(\mathbf{u}+\mathbf{h})\} - E\{Z_v(\mathbf{u})\} E\{Z_\bullet(\mathbf{u}+\mathbf{h})\} = \frac{1}{|v|} \sum_{v(\mathbf{u})} C_{\bullet\bullet}(\mathbf{h}') \tag{12}
$$

After establishing the correlations between different scales, we can use them to account for the different volume supports of the conditioning data in the multiscale modeling using a direct kriging or direct sequential simulation framework.

Theory of Exact Downscaling with Direct Kriging

Assume a data domain *A* is stationary. Let $\{z_\bullet(\mathbf{u}'_i), i = 1,\ldots n, \quad \forall\, \mathbf{u}'_i \in A\}$ be the available point data, and let $\{z_V(\mathbf{u}''_j), j = 1,\ldots n', \quad \forall\, \mathbf{u}''_j \in A\}$ be the block data that is exhaustively available over *A*. Apply simple kriging, we can using the block and point data to estimate the values at the model scale:

$$
Z_v^*(\mathbf{u}) - m = \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u})(Z_\bullet(\mathbf{u}'_i) - m) + \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u})\,(Z_V(\mathbf{u}''_{i'}) - m) \qquad \text{or}
$$

$$Z_v^*(\mathbf{u}) = \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}) Z_\bullet(\mathbf{u}'_i) + \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u}) Z_V(\mathbf{u}''_{i'}) + \left[1 - \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}) - \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u})\right] m \tag{13}$$

where $\lambda_{V i'}(\mathbf{u})$ is the kriging weight for block data $Z_V(\mathbf{u}'')$, and $\lambda_\bullet$ is the weight for point data $Z_\bullet(\mathbf{u}')$. More precisely, this equation is *simple block cokriging*.

By minimizing the error variance, we can obtain the simple block cokriging system:

$$\begin{cases} \sum_{i'=1}^{n'} \lambda_{V i'} \overline{C}_{V_{i'} V_j} + \sum_{i=1}^{n} \lambda_{\bullet i} \overline{C}_{\bullet_i V_j} = \overline{C}_{v(\mathbf{u}) V_j} & j = 1, \ldots m \\ \sum_{i'=1}^{n'} \lambda_{V i'} \overline{C}_{V_{j'} \bullet_i} + \sum_{i=1}^{n} \lambda_{\bullet i} \overline{C}_{\bullet_{j'} \bullet_i} = \overline{C}_{v(\mathbf{u}) \bullet_{j'}} & j' = 1, \ldots m' \end{cases} \tag{14}$$

And the minimized simple block cokriging variance can be expressed as below:

$$\sigma^2_{sbck}(\mathbf{u}) = \sigma_v^2 - \sum_{i=1}^{n} \lambda_{\bullet i} \overline{C}_{\bullet_i v(\mathbf{u})} - \sum_{i'=1}^{n'} \lambda_{V i'} \overline{C}_{V_{i'} v(\mathbf{u})} \tag{15}$$

where σ_v^2 is the variance at the model scale.

In the downscaling process, the block is expected to be reproduced exactly so that the fine scale model is consistent with the large scale model and no bias is introduced by the scaling process. A theorem on the exact reproduction of block data using the direct simple block cokriging is proposed. Then a proof is given.

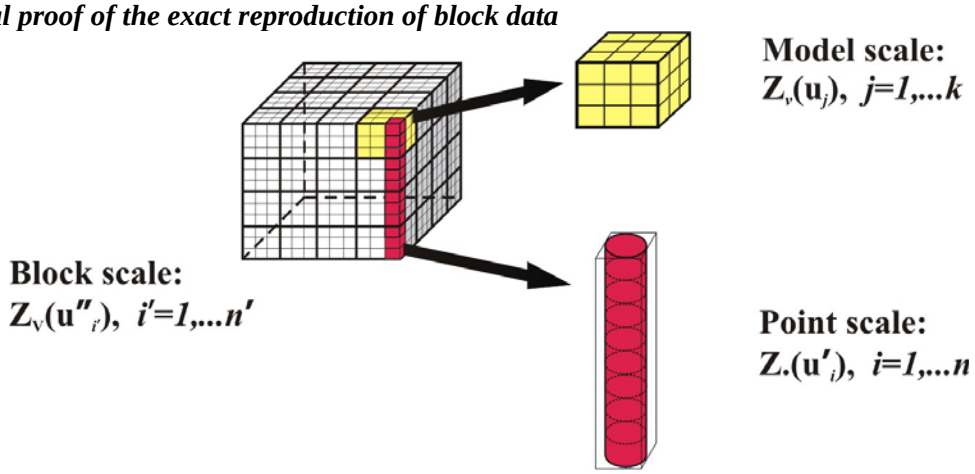
Theorem I: simple block cokriging with block data and point data is an exact downscaling method:

$$Z_v^*(\mathbf{u}) = \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}) Z_\bullet(\mathbf{u}'_i) + \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u}) Z_V(\mathbf{u}''_{i'}) + \left[1 - \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}) - \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u})\right] m$$

That is, the estimated values of the small cells in a block can exactly reproduce the block value:

$$Z_V^*(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} Z^*(\mathbf{u}_i) = Z_V(\mathbf{u}) \tag{16}$$

General proof of the exact reproduction of block data



Consider a block constituting of k small cells, and there are n point data available. Then, for each small cell at $\mathbf{u}_j$, ($j=1 \ldots k$), the kriging estimates are:

$$Z_v^*(\mathbf{u}_j) = \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}_j) Z_{\bullet}(\mathbf{u}'_i) + \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u}_j) Z_V(\mathbf{u}''_{i'}) + \left[1 - \sum_{i=1}^{n} \lambda_{\bullet i}(\mathbf{u}_j) - \sum_{i'=1}^{n'} \lambda_{V i'}(\mathbf{u}_j) \right] m$$

and the kriging system in matrix format is

$$\begin{pmatrix}
\overline{C}_{V_1V_1} & \overline{C}_{V_1V_2} & \cdots & \overline{C}_{V_1V_{n'}} & \overline{C}_{V_1\bullet_1} & \overline{C}_{V_1\bullet_2} & \cdots & \overline{C}_{V_1\bullet_n} \\
\overline{C}_{V_2V_1} & \overline{C}_{V_2V_2} & \cdots & \overline{C}_{V_2V_{n'}} & \overline{C}_{V_2\bullet_1} & \overline{C}_{V_2\bullet_2} & \cdots & \overline{C}_{V_2\bullet_n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\overline{C}_{V_{n'}V_1} & \overline{C}_{V_{n'}V_2} & \cdots & \overline{C}_{V_{n'}V_{n'}} & \overline{C}_{V_{n'}\bullet_1} & \overline{C}_{V_{n'}\bullet_2} & \cdots & \overline{C}_{V_{n'}\bullet_n} \\
\overline{C}_{\bullet_1V_1} & \overline{C}_{\bullet_1V_2} & \cdots & \overline{C}_{\bullet_1V_{n'}} & C_{\bullet_1\bullet_1} & C_{\bullet_1\bullet_2} & \cdots & C_{\bullet_1\bullet_n} \\
\overline{C}_{\bullet_2V_1} & \overline{C}_{\bullet_2V_2} & \cdots & \overline{C}_{\bullet_2V_{n'}} & C_{\bullet_2\bullet_1} & C_{\bullet_2\bullet_2} & \cdots & C_{\bullet_2\bullet_n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\overline{C}_{\bullet_nV_1} & \overline{C}_{\bullet_nV_2} & \cdots & \overline{C}_{\bullet_nV_{n'}} & C_{\bullet_n\bullet_1} & C_{\bullet_n\bullet_2} & \cdots & C_{\bullet_n\bullet_n}
\end{pmatrix}
\begin{pmatrix}
\lambda_{V_1}(\mathbf{u}_j) \\
\lambda_{V_2}(\mathbf{u}_j) \\
\vdots \\
\lambda_{V_{n'}}(\mathbf{u}_j) \\
\lambda_{\bullet_1}(\mathbf{u}_j) \\
\lambda_{\bullet_2}(\mathbf{u}_j) \\
\vdots \\
\lambda_{\bullet_n}(\mathbf{u}_j)
\end{pmatrix}
=
\begin{pmatrix}
\overline{C}_{V_1v(\mathbf{u}_j)} \\
\overline{C}_{V_2v(\mathbf{u}_j)} \\
\vdots \\
\overline{C}_{V_{n'}v(\mathbf{u}_j)} \\
\overline{C}_{\bullet_1v(\mathbf{u}_j)} \\
\overline{C}_{\bullet_2v(\mathbf{u}_j)} \\
\vdots \\
\overline{C}_{\bullet_nv(\mathbf{u}_j)}
\end{pmatrix}$$

If we summarize all the k kriging systems together, we get

$$\begin{pmatrix}
\overline{C}_{V_1V_1} & \overline{C}_{V_1V_2} & \cdots & \overline{C}_{V_1V_{n'}} & \overline{C}_{V_1\bullet_1} & \overline{C}_{V_1\bullet_2} & \cdots & \overline{C}_{V_1\bullet_n} \\
\overline{C}_{V_2V_1} & \overline{C}_{V_2V_2} & \cdots & \overline{C}_{V_2V_{n'}} & \overline{C}_{V_2\bullet_1} & \overline{C}_{V_2\bullet_2} & \cdots & \overline{C}_{V_2\bullet_n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\overline{C}_{V_{n'}V_1} & \overline{C}_{V_{n'}V_2} & \cdots & \overline{C}_{V_{n'}V_{n'}} & \overline{C}_{V_{n'}\bullet_1} & \overline{C}_{V_{n'}\bullet_2} & \cdots & \overline{C}_{V_{n'}\bullet_n} \\
\overline{C}_{\bullet_1V_1} & \overline{C}_{\bullet_1V_2} & \cdots & \overline{C}_{\bullet_1V_{n'}} & C_{\bullet_1\bullet_1} & C_{\bullet_1\bullet_2} & \cdots & C_{\bullet_1\bullet_n} \\
\overline{C}_{\bullet_2V_1} & \overline{C}_{\bullet_2V_2} & \cdots & \overline{C}_{\bullet_2V_{n'}} & C_{\bullet_2\bullet_1} & C_{\bullet_2\bullet_2} & \cdots & C_{\bullet_2\bullet_n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\overline{C}_{\bullet_nV_1} & \overline{C}_{\bullet_nV_2} & \cdots & \overline{C}_{\bullet_nV_{n'}} & C_{\bullet_n\bullet_1} & C_{\bullet_n\bullet_2} & \cdots & C_{\bullet_n\bullet_n}
\end{pmatrix}
\begin{pmatrix}
\sum_{j=1}^{k} \lambda_{V_1}(\mathbf{u}_j) \\
\sum_{j=1}^{k} \lambda_{V_2}(\mathbf{u}_j) \\
\vdots \\
\sum_{j=1}^{k} \lambda_{V_{n'}}(\mathbf{u}_j) \\
\sum_{j=1}^{k} \lambda_{\bullet_1}(\mathbf{u}_j) \\
\sum_{j=1}^{k} \lambda_{\bullet_2}(\mathbf{u}_j) \\
\vdots \\
\sum_{j=1}^{k} \lambda_{\bullet_n}(\mathbf{u}_j)
\end{pmatrix}
=
\begin{pmatrix}
\sum_{j=1}^{k} \overline{C}_{V_1v_j} \\
\sum_{j=1}^{k} \overline{C}_{V_2v_j} \\
\vdots \\
\sum_{j=1}^{k} \overline{C}_{V_{n'}v_j} \\
\sum_{j=1}^{k} \overline{C}_{\bullet_1v_j} \\
\sum_{j=1}^{k} \overline{C}_{\bullet_2v_j} \\
\vdots \\
\sum_{j=1}^{k} \overline{C}_{\bullet_nv_j}
\end{pmatrix}$$

From the volume averaged covariance relationship, we know that:

$$\overline{C}_{V_iV_j} = \frac{1}{k} \sum_{j=1}^{k} \overline{C}_{V_iv_j} \qquad i = 1, 2, \ldots n'$$

$$\overline{C}_{\bullet_iV_j} = \frac{1}{k} \sum_{j=1}^{k} \overline{C}_{\bullet_iv_j} \qquad i = 1, 2, \ldots n$$

Then, the matrix becomes:

$$\begin{pmatrix} \overline{C}_{V_1V_1} & \overline{C}_{V_1V_2} & \cdots & \overline{C}_{V_1V_{n'}} & \overline{C}_{V_1\bullet_1} & \overline{C}_{V_1\bullet_2} & \cdots & \overline{C}_{V_1\bullet_n} \\ \overline{C}_{V_2V_1} & \overline{C}_{V_2V_2} & \cdots & \overline{C}_{V_2V_{n'}} & \overline{C}_{V_2\bullet_1} & \overline{C}_{V_2\bullet_2} & \cdots & \overline{C}_{V_2\bullet_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{V_{n'}V_1} & \overline{C}_{V_{n'}V_2} & \cdots & \overline{C}_{V_{n'}V_{n'}} & \overline{C}_{V_{n'}\bullet_1} & \overline{C}_{V_{n'}\bullet_2} & \cdots & \overline{C}_{V_{n'}\bullet_n} \\ \overline{C}_{\bullet_1V_1} & \overline{C}_{\bullet_1V_2} & \cdots & \overline{C}_{\bullet_1V_{n'}} & C_{\bullet_1\bullet_1} & C_{\bullet_1\bullet_2} & \cdots & C_{\bullet_1\bullet_n} \\ \overline{C}_{\bullet_2V_1} & \overline{C}_{\bullet_2V_2} & \cdots & \overline{C}_{\bullet_2V_{n'}} & C_{\bullet_2\bullet_1} & C_{\bullet_2\bullet_2} & \cdots & C_{\bullet_2\bullet_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{\bullet_nV_1} & \overline{C}_{\bullet_nV_2} & \cdots & \overline{C}_{\bullet_nV_{n'}} & C_{\bullet_n\bullet_1} & C_{\bullet_n\bullet_2} & \cdots & C_{\bullet_n\bullet_n} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{k} \lambda_{V_1}(\mathbf{u}_j) \\ \sum_{j=1}^{k} \lambda_{V_2}(\mathbf{u}_j) \\ \vdots \\ \sum_{j=1}^{k} \lambda_{V_{n'}}(\mathbf{u}_j) \\ \sum_{j=1}^{k} \lambda_{\bullet_1}(\mathbf{u}_j) \\ \sum_{j=1}^{k} \lambda_{\bullet_2}(\mathbf{u}_j) \\ \vdots \\ \sum_{j=1}^{k} \lambda_{\bullet_n}(\mathbf{u}_j) \end{pmatrix} = \begin{pmatrix} k\overline{C}_{V_1V_1} \\ k\overline{C}_{V_2V_1} \\ \vdots \\ k\overline{C}_{V_kV_1} \\ k\overline{C}_{\bullet_1V_1} \\ k\overline{C}_{\bullet_2V_1} \\ \vdots \\ k\overline{C}_{\bullet_nV_1} \end{pmatrix}$$

The right hand side is k multiply the first column of the covariance matrix. Thus, one solution is:

$$\begin{cases} \sum_{j=1}^{k} \lambda_{V_1} = k \\ \sum_{j=1}^{k} \lambda_{V_2} = \sum_{j=1}^{k} \lambda_{V_3} = \cdots = \sum_{j=1}^{k} \lambda_{V_{n'}} = 0 \\ \sum_{j=1}^{k} \lambda_{\bullet_1} = \sum_{j=1}^{k} \lambda_{\bullet_2} = \cdots = \sum_{j=1}^{k} \lambda_{\bullet_n} = 0 \end{cases} \tag{5-17}$$

Because it is a kriging system, this solution is unique. This solution indicates that the sum of the weights of the collocated block in all estimates in the block is the number of cells in the block, and the sum of the weights of point data and other block data in all estimates is zero.

Using this solution, the block value calculated from all the estimates is:

$$\begin{aligned} z_V^* &= \frac{1}{k}\sum_{j=1}^{k} z_v^*(\mathbf{u}_j) \\ &= \frac{1}{k}\sum_{j=1}^{k}\left\{\sum_{i=1}^{n}\lambda_{\bullet i}(\mathbf{u}_j)Z_\bullet(\mathbf{u}'_i) + \sum_{i'=1}^{n'}\lambda_{Vi'}(\mathbf{u}_j)Z_V(\mathbf{u}''_{i'}) + \left[1 - \sum_{i=1}^{n}\lambda_{\bullet i}(\mathbf{u}_j) - \sum_{i'=1}^{n'}\lambda_{Vi'}(\mathbf{u}_j)\right]m\right\} \\ &= \frac{1}{k}\left\{\sum_{j=1}^{k}\lambda_{V1}(\mathbf{u}_j)Z_{V1} + k\,m - \sum_{j=1}^{k}\lambda_{V1}(\mathbf{u}_j)m\right\} \\ &= Z_V \end{aligned}$$

Therefore, the block data can always be reproduced exactly with the simple block cokriging. And this simple block cokriging is named *exact downscaling with direct kriging* (**EDDK**).

According to the characteristics of kriging, the kriging results are smooth with less extreme values. The missing variance is the kriging variance. To generate the correct realizations, we need to use the direct sequential simulation (DSS) approach.

Theory of Exact Downscaling with Direct Sequential Simulation

Assume a data domain *A* is stationary. Let $\{z_\bullet(\mathbf{u}'_i), i=1,...n, \quad \forall \mathbf{u}'_i \in A\}$ be the available point data, and let $\{z_V(\mathbf{u}''_j), j=1,...n', \quad \forall \mathbf{u}''_j \in A\}$ be the block data that is exhaustively available over A. Apply direct sequential simulation, the block, point data and previously simulated data are used to simulate the values at the model scale:

$$Z_{vs}(\mathbf{u}) = \lambda_V Z_V(\mathbf{u}') + \sum_{i=1}^{n} \lambda_{vi} Z_{vs}(\mathbf{u}_i) + \sum_{i'=1}^{n} \lambda_{\bullet i'} Z_\bullet(\mathbf{u}''_{i'}) + \left[1 - \lambda_V - \sum_{i=1}^{n} \lambda_{vi} - \sum_{i'=1}^{n'} \lambda_{\bullet i'}\right] m + R(\mathbf{u}) \tag{18}$$

where $Z_V(\mathbf{u}')$ is the collocated block value, only the collocated block datum is used. And $Z_\bullet(\mathbf{u}'')$ is the point data, $Z_{vs}(\mathbf{u}_l)$ is the previously simulated data, $R(\mathbf{u})$ is the random residual at the simulated location $\mathbf{u}$.

The kriging system is written as:

$$\begin{cases} \lambda_V \overline{C}_{VV} + \sum_{i=1}^{n} \lambda_{vi} \overline{C}_{Vv_i} + \sum_{i'=1}^{n'} \lambda_{\bullet i'} \overline{C}_{V\bullet_{i'}} = \overline{C}_{Vv(\mathbf{u})} \\ \lambda_V \overline{C}_{Vv_j} + \sum_{i=1}^{n} \lambda_{vi} \overline{C}_{v_i v_j} + \sum_{i'=1}^{n'} \lambda_{\bullet i'} \overline{C}_{\bullet_{i'} v_j} = \overline{C}_{v_j v(\mathbf{u})}, \qquad j = 1,..n \\ \lambda_V \overline{C}_{V\bullet_{j'}} + \sum_{i=1}^{n} \lambda_{vi} \overline{C}_{v_i \bullet_{j'}} + \sum_{i'=1}^{n'} \lambda_{\bullet i'} C_{\bullet i' \bullet j'} = \overline{C}_{\bullet_{j'} v(\mathbf{u})}, \qquad j' = 1,..n' \end{cases} \tag{19}$$

And the simple block cokriging variance is

$$\sigma^2_{sbck}(\mathbf{u}) = \sigma^2_v - \lambda_V \overline{C}_{Vv(\mathbf{u})} - \sum_{i=1}^{n} \lambda_{vi} \overline{C}_{v_i v(\mathbf{u})} - \sum_{i'=1}^{n'} \lambda_{\bullet i'} \overline{C}_{\bullet_{i'} v(\mathbf{u})} \tag{20}$$

where σ_v^2 is the variance at model scale.

Theorem II: Direct sequential simulation with block data and point data is an exact downscaling method**:**

$$Z_{vs}(\mathbf{u}) = \lambda_V Z_V(\mathbf{u}') + \sum_{i=1}^{n} \lambda_{vi} Z_{vs}(\mathbf{u}_i) + \sum_{i'=1}^{n} \lambda_{\bullet i'} Z_\bullet(\mathbf{u}''_{i'}) + \left[1 - \lambda_V - \sum_{i=1}^{n} \lambda_{vi} - \sum_{i'=1}^{n'} \lambda_{\bullet i'}\right] m + R(\mathbf{u})$$

that is, the simulated values of the small cells in a block can exactly reproduce the block value:

$$Z_{Vs}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} Z_{vs}(\mathbf{u}'_i) = Z_V(\mathbf{u}) \tag{21}$$

General Proof of Theorem II:

Consider a block constituting of *k* small cells, and there are *n* point data are available. Perform the DSS to simulate the cells in the block following a random path; the previously simulated cells are

added into the conditioning data for simulating the next cell, until reach the last cell located at $\mathbf{u}_k$, the simulated value is:

$$Z_{vs}(\mathbf{u}_k) = \lambda_V Z_V(\mathbf{u}') + \sum_{i=1}^{n} \lambda_{vi} Z_{vs}(\mathbf{u}_i) + \sum_{i'=1}^{n'} \lambda_{\bullet i'} Z_{\bullet}(\mathbf{u}''_{i'}) + \left[1 - \lambda_V - \sum_{i=1}^{n} \lambda_{vi} - \sum_{i'=1}^{n'} \lambda_{\bullet i'}\right] m + R(\mathbf{u}_k)$$

where among the n previously simulated cells, there are $k-1$ of them inside the block. And the kriging system in matrix format:

$$\begin{pmatrix} \overline{C}_{VV} & \overline{C}_{Vv_1} & \cdots & \overline{C}_{Vv_n} & \overline{C}_{V\bullet_1} & \overline{C}_{V\bullet_2} & \cdots & \overline{C}_{V\bullet_{n'}} \\ \overline{C}_{v_1V} & \overline{C}_{v_1v_1} & \cdots & \overline{C}_{v_1v_n} & \overline{C}_{v_1\bullet_1} & \overline{C}_{v_2\bullet_2} & \cdots & \overline{C}_{v_2\bullet_{n'}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{v_nV} & \overline{C}_{v_nv_1} & \cdots & \overline{C}_{v_nv_n} & \overline{C}_{v_n\bullet_1} & \overline{C}_{v_n\bullet_2} & \cdots & \overline{C}_{v_n\bullet_{n'}} \\ \overline{C}_{\bullet_1V} & \overline{C}_{\bullet_1v_1} & \cdots & \overline{C}_{\bullet_1v_n} & C_{\bullet_1\bullet_1} & C_{\bullet_1\bullet_2} & \cdots & C_{\bullet_1\bullet_{n'}} \\ \overline{C}_{\bullet_2V} & \overline{C}_{\bullet_2v_1} & \cdots & \overline{C}_{\bullet_2v_n} & C_{\bullet_2\bullet_1} & C_{\bullet_2\bullet_2} & \cdots & C_{\bullet_2\bullet_{n'}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{\bullet_{n'}V} & \overline{C}_{\bullet_{n'}v_1} & \cdots & \overline{C}_{\bullet_{n'}v_n} & C_{\bullet_{n'}\bullet_1} & C_{\bullet_{n'}\bullet_2} & \cdots & C_{\bullet_{n'}\bullet_{n'}} \end{pmatrix} \begin{pmatrix} \lambda_V(\mathbf{u}_k) \\ \lambda_{v_1}(\mathbf{u}_k) \\ \vdots \\ \lambda_{v_n}(\mathbf{u}_k) \\ \lambda_{\bullet_1}(\mathbf{u}_k) \\ \lambda_{\bullet_2}(\mathbf{u}_k) \\ \vdots \\ \lambda_{\bullet_{n'}}(\mathbf{u}_k) \end{pmatrix} = \begin{pmatrix} \overline{C}_{Vv(\mathbf{u}_k)} \\ \overline{C}_{v_1v(\mathbf{u}_k)} \\ \vdots \\ \overline{C}_{v_nv(\mathbf{u}_k)} \\ \overline{C}_{\bullet_1v(\mathbf{u}_k)} \\ \overline{C}_{\bullet_2v(\mathbf{u}_k)} \\ \vdots \\ \overline{C}_{\bullet_{n'}v(\mathbf{u}_k)} \end{pmatrix}$$

Let's solve it. From the volume averaged correlations, we can get:

$$\overline{C}_{VV} = \frac{1}{k}\sum_{i=1}^{k} \overline{C}_{Vv_i} \qquad\qquad k\overline{C}_{VV} - \sum_{i=1}^{k-1} \overline{C}_{Vv_i} = \overline{C}_{Vv(\mathbf{u}_k)}$$

or

$$\overline{C}_{Vv_j} = \frac{1}{k}\sum_{i=1}^{k} \overline{C}_{v_jv_i}, \quad j = 1,..n \quad \Rightarrow \quad k\overline{C}_{Vv_j} - \sum_{i=1}^{k-1} \overline{C}_{v_iv_j} = \overline{C}_{v_jv(\mathbf{u}_k)}, \quad j = 1,..n$$

$$\overline{C}_{V\bullet_{j'}} = \frac{1}{k}\sum_{i=1}^{k} C_{v_i\bullet_{j'}}, \quad j' = 1,..n' \quad \Rightarrow \quad k\overline{C}_{V\bullet_{j'}} - \sum_{i=1}^{k-1} \overline{C}_{v_i\bullet_{j'}} = \overline{C}_{v_n\bullet_{j'}}, \quad j' = 1,..n'$$

Then, the kriging system becomes:

$$\begin{pmatrix} \overline{C}_{VV} & \overline{C}_{Vv_1} & \cdots & \overline{C}_{Vv_n} & \overline{C}_{V\bullet_1} & \overline{C}_{V\bullet_2} & \cdots & \overline{C}_{V\bullet_{n'}} \\ \overline{C}_{v_1V} & \overline{C}_{v_1v_1} & \cdots & \overline{C}_{v_1v_n} & \overline{C}_{v_1\bullet_1} & \overline{C}_{v_2\bullet_2} & \cdots & \overline{C}_{v_2\bullet_{n'}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{v_nV} & \overline{C}_{v_nv_1} & \cdots & \overline{C}_{v_nv_n} & \overline{C}_{v_n\bullet_1} & \overline{C}_{v_n\bullet_2} & \cdots & \overline{C}_{v_n\bullet_{n'}} \\ \overline{C}_{\bullet_1V} & \overline{C}_{\bullet_1v_1} & \cdots & \overline{C}_{\bullet_1v_n} & C_{\bullet_1\bullet_1} & C_{\bullet_1\bullet_2} & \cdots & C_{\bullet_1\bullet_{n'}} \\ \overline{C}_{\bullet_2V} & \overline{C}_{\bullet_2v_1} & \cdots & \overline{C}_{\bullet_2v_n} & C_{\bullet_2\bullet_1} & C_{\bullet_2\bullet_2} & \cdots & C_{\bullet_2\bullet_{n'}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \overline{C}_{\bullet_{n'}V} & \overline{C}_{\bullet_{n'}v_1} & \cdots & \overline{C}_{\bullet_{n'}v_n} & C_{\bullet_{n'}\bullet_1} & C_{\bullet_{n'}\bullet_2} & \cdots & C_{\bullet_{n'}\bullet_{n'}} \end{pmatrix} \begin{pmatrix} \lambda_V(\mathbf{u}_k) \\ \lambda_{v_1}(\mathbf{u}_k) \\ \vdots \\ \lambda_{v_n}(\mathbf{u}_k) \\ \lambda_{\bullet_1}(\mathbf{u}_k) \\ \lambda_{\bullet_2}(\mathbf{u}_k) \\ \vdots \\ \lambda_{\bullet_{n'}}(\mathbf{u}_k) \end{pmatrix} = \begin{pmatrix} k\overline{C}_{VV} - \sum_{i=1}^{k-1} \overline{C}_{Vv_i} \\ k\overline{C}_{Vv_1} - \sum_{i=1}^{k-1} \overline{C}_{v_iv_1} \\ \vdots \\ k\overline{C}_{Vv_n} - \sum_{i=1}^{k-1} \overline{C}_{v_iv_{k-1}} \\ k\overline{C}_{V\bullet_1} - \sum_{i=1}^{k-1} \overline{C}_{v_i\bullet_1} \\ k\overline{C}_{V\bullet_2} - \sum_{i=1}^{k-1} \overline{C}_{v_i\bullet_2} \\ \vdots \\ k\overline{C}_{V\bullet_{n'}} - \sum_{i=1}^{k-1} \overline{C}_{v_i\bullet_{n'}} \end{pmatrix}$$

This matrix gives a unique solution:

$$\begin{cases} \lambda_V = k \\ \lambda_{vi} = -1, & i = 1, .. k-1 \\ \lambda_{vi} = 0, & i = k, .. n \\ \lambda_{\bullet i'} = 0, & i' = 1, .. n' \end{cases} \tag{22}$$

Now, let's check the variance:

$$\sigma^2_{sbck}(\mathbf{u}_k) = \sigma^2_v - \lambda_V \overline{C}_{Vv(\mathbf{u})} - \sum_{i=1}^{n} \lambda_{vi} \overline{C}_{v_i v(\mathbf{u})} - \sum_{i'=1}^{n'} \lambda_{\bullet i'} \overline{C}_{\bullet_{i'} v(\mathbf{u})} = \overline{C}_{v_k v_k} - k\overline{C}_{Vv(\mathbf{u})} + \sum_{i=1}^{k-1} \overline{C}_{v_i v_k}$$

$$= \sum_{i=1}^{k} \overline{C}_{v_i v_k} - k\overline{C}_{Vv(\mathbf{u})} = 0 \tag{23}$$

Thus, the random residual at the last cell $R(\mathbf{u}_k) \equiv 0$ because it follows a distribution with zero mean and zero variance. Then, the simulated value at the last location is

$$Z_{vs}(\mathbf{u}_k) = k Z_V - \sum_{i=1}^{k-1} Z_{vs}(\mathbf{u}_i) \tag{24}$$

This equation indicates that the last simulated value is actually the block value multiply the number of model cells in the block and subtracts all the previously simulated cells in the block. Therefore, the block average of simulated values at all locations in the block is

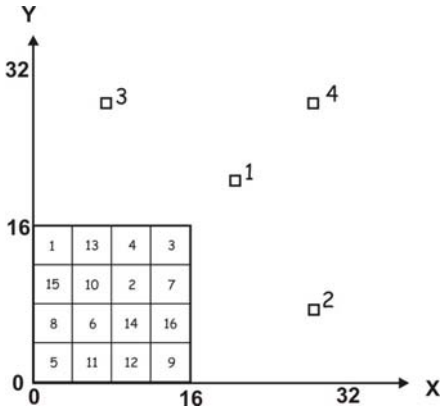
$$Z_{Vs} = \frac{1}{k}\left[\sum_{i=1}^{k-1} Z_{vs}(\mathbf{u}_i) + Z_{vs}(\mathbf{u}_k)\right] = \frac{1}{k}\left[\sum_{i=1}^{k-1} Z_{vs}(\mathbf{u}_i) + k Z_V - \sum_{i=1}^{k-1} Z_{vs}(\mathbf{u}_i)\right] = Z_V$$

Thus, the block data can always be exactly reproduces in direct sequential simulation. And this exact downscaling technique is named *exact downscaling with direct sequential simulation* (**EDDSS**).

Some interesting aspects of EDDSS

Because the previously simulated cells are used in EDDSS, only the collocated block is used for computer efficiency. Kriging weights of block and point data will be affected by the previously simulated data.

Consider downscaling a block at scale of 16 m by 16 m into 16 small cells at scale of 4 m by 4 m as shown in the figure below. There are also four point data at a scale of 1 m by 1 m close to the block.



1	13	4	3
15	10	2	7
8	6	14	16
5	11	12	9

EDDSS is performed to use the block and point data. The results are shown in the table below. The 16 small cells are simulated with a random path. The weights of data in each cell are given in the 16 columns. The kriged estimate and the variance are listed in the last two rows. The table shows the block value is exactly reproduced.

Data					Kriging Weights in Simulated Cell															Summary
type	x	y	z	value	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	/Average
b	8	8	0.5	**48.10**	0.88	1.23	0.03	-0.20	2.21	1.49	1.33	0.13	4.68	1.80	3.55	3.76	2.86	5.58	8.10	16.00	53.43
p	20.5	20.5	0.5	58	-0.04	0.01	0.32	-0.03	0.05	0.01	-0.03	0.01	0.00	-0.01	0.01	0.01	0.00	-0.01	0.01	0.00	0.30
p	28.5	7.5	0.5	28	0.00	-0.04	0.04	-0.03	-0.02	-0.04	0.03	0.00	0.03	0.02	0.00	-0.02	0.00	-0.01	-0.01	0.00	-0.04
p	7.5	28.5	0.5	48	0.09	-0.03	0.05	0.04	0.01	-0.02	-0.01	-0.02	0.02	-0.02	-0.01	0.00	0.02	0.01	-0.01	0.00	0.12
p	28.5	28.5	0.5	38	0.03	-0.02	-0.06	0.00	0.00	-0.02	0.00	-0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.05
s	2	10	0.5																	-1.00	-1.00
s	10	6	0.5																-0.74	-1.00	-1.74
s	6	14	0.5															-0.28	-0.59	-1.00	-1.88
s	10	2	0.5														-0.25	-0.11	-0.42	-1.00	-1.77
s	6	2	0.5													0.23	-0.12	-0.43	-0.51	-1.00	-1.82
s	6	10	0.5												-0.48	-0.42	0.11	-0.58	-0.27	-1.00	-2.64
s	14	2	0.5											-0.32	-0.37	-0.03	-0.28	-0.49	-0.74	-1.00	-3.23
s	2	6	0.5										-0.41	0.04	-0.46	-0.35	-0.34	-0.61	-0.24	-1.00	-3.37
s	14	10	0.5									-0.07	-0.23	-0.30	-0.42	-0.46	-0.31	-0.50	-0.78	-1.00	-4.07
s	6	6	0.5								-0.40	0.42	-0.81	0.02	0.10	-0.27	-0.34	-0.02	-0.52	-1.00	-2.82
s	2	2	0.5							0.08	-0.19	0.44	-0.38	-0.30	0.21	-0.31	-0.14	-0.30	-0.57	-1.00	-2.45
s	10	14	0.5						-0.07	-0.09	-0.38	-0.03	-0.48	-0.06	-0.27	-0.33	0.35	-0.42	-0.51	-1.00	-3.28
s	14	14	0.5					0.50	-0.23	-0.27	0.48	0.01	-0.23	-0.11	-0.16	-0.16	-0.26	-0.32	-0.44	-1.00	-2.20
s	10	10	0.5				0.63	0.47	-0.75	0.02	0.36	-0.08	-0.64	0.24	-0.33	-0.24	-0.37	-0.02	-0.49	-1.00	-2.19
s	2	14	0.5			-0.11	-0.05	0.25	-0.23	-0.12	-0.22	0.21	-0.61	0.00	-0.39	-0.43	0.07	-0.50	-0.27	-1.00	-3.40
Kriging Estimate					47.41	49.29	51.86	50.35	47.30	48.13	48.84	47.17	46.45	48.25	47.76	48.08	48.43	48.42	46.75	45.06	**48.10**
Kriging Variance					0.32	0.12	0.21	0.09	0.19	0.08	0.07	0.09	0.08	0.05	0.05	0.04	0.04	0.02	0.02	0.00	

The kriging weights accounts for the data closeness, redundancy, and support volume. The weight of the block datum is not always higher than the weights of the point data. At the cells 3, 4, and 8, the weight of the block datum is small or even negative due to the screen effect of adjacent simulated cells. The weights of the point data are always low due to their small volume and long distance from the simulated cell. When simulating the last few cells (14, 15 and 16 in this example), the weight of the block datum is going up fast, while the weights of the point data become close to zero, and the weights of previously simulated cells become negative. In the last cell, the weight of block datum is the discretization number; the weights of the point data are zero; and the weights of the previously simulated cells are all -1. And following the simulation sequence, the kriging variance is going down, at the last cell, the variance is zero. This matches what is shown in the proof of Theorem II.

Implementation details

Global mean

The block data mean should be used as the global mean because the block data is exhaustively over the whole model area. It is normal that the mean of sampled point data is different from block data mean. Declustering of point data may be necessary to reduce the difference.

Discretization level

Large discretization number between block scale and point scale can cause a very high computation time and low precision of simulated results. If downscale a 3-D model in three directions with a discretization number of 4, it is 4,096 times point scale covariance calculations for the block covariance when the model scale is same as the point scale. If the discretization number between the model scale and point scale is 4 in three directions, 16,777,216 times point scale covariance calculations are needed. Such large number could cause round-up error and resulting the block data are not exactly reproduced. Increasing data precision will need more computering space and time. Covariance lookup tables for all three scales are necessary to reduce the computation time.

The EDDSS requires all the previously simulated cells in the block to be used. The kriging system includes the collocated block datum, the nearby point data and the nearby previously simulated cells. A large discretization number will give a large kriging system. Solving a large system requires more computation cost. However, it is wise to downscale small blocks and to avoid downscaling in three directions.

Last Cell Correction

It is interesting in EDDSS that the block datum starts to take control and screens out the previously simulated cells in the last few cells, and forces the block data to be exactly reproduced at the last cell. The good side of the block value control is the block data are exact reproduced; the bad side is the simulated values of last cells are often out of the valid data range. Some post corrections or re-simulation in such blocks are almost unavoidable. Actually, an invalid value in the last cell indicates an error in the values of the previously simulated cells. Re-simulate all the cells in the block with a new random path would be more appropriate than directly modifying the simulated values of those cells. The number of re-simulation times should be restricted to not significantly increase the computation time.

Simulation Path

Two simulation paths could be used for the EDDSS. One is a random path for all the cells over the entire model area. However, this path is loose with the local block data control, will result more invalid values in the last cells. It also makes the corrections or re-simulation of the block difficult to apply. Another simulation path is signing a random path to each block, and simulated all the cells in each block following a random path. The advantage of this path is easy to re-simulate the block when it is necessary.

Search Strategies

No need for block data search because only use the collocated block; using the super block search for point data because well (point) data are normally non-gridded data; using the spiral search for the previously simulated nodes because they are regularly grids. There are two parts of previously simulated nodes: inside of block and outside of block:

Inside block: all the simulated nodes need to be included.

Outside block: a maximum number of nodes need to be specified. And the octant search is needed to restrict the number of nodes from one nearby block. This is because all of the nodes in one closest block can fill the maximum number.

Histogram Reproduction and Proportion Effect

When working in Gaussian space, the kriging estimate and variance are sufficient to characterize the local distributions of uncertainty, because the local distributions are Gaussian. And the back transform using the global distribution ensures the histogram reproduction. In direct space, the kriging estimate and the estimation variance are not adequately to describe the local distributions because the shapes of the distributions are not known. Deutsch *et al.* (2002) proposed a method to use the global distribution to build CCDF lookup tables for the local distributions. A series of quantiles of local Gaussian distributions are back transformed into direct space to characterize the local distributions.

$$Z_L(q) = F^{-1}[G_{\{0,1\}}[G^{-1}{}_{\{m,\sigma\}}(q)]]$$

Then, the mean and variance of each distribution are calculated for index to look up with the kriging estimate and variance. The graphical representation of the transformations applied to calculate the local distributions of uncertainty is shown in Figure 4. An implicit assumption of the CCDF lookup method is that all the quantiles of normal distributions follows a standard normal distribution, which is zero mean, unit variance and more than 99% values between -3 and 3. Only a mean limit of 1.5 and variance limit of 1 can reach the standard normal distribution. As an example, the results of using the CCDF lookup-table method are shown in Figure 5. The global distribution (middle left) is a lognormal distribution. The first row shows the cross plots of mean vs. variance in CCDF table (left) and in kriging results (right). The CCDF tables are not able to cover all the points from the kriging. The means of the CCDF tables are less than 60 for variance under 1000. But lots of kriging estimates are higher than 60. Thus the matching results give a flat line around 60 (bottom right).

One possible reason for the mismatch is the kriging variance is not correct. The kriging variance is independent from the kriging mean. The real data always show a relation between the mean and variance, which is called proportion effect. This relation is usually calculated by using linear regression with results from moving window method. Correcting the kriging variance using the fitted function is problematic. One problem is the relation function changes with the windows size. The local uncertainty distribution with kriging mean and variance is not the distribution of search neighborhood. So it is hard to find the scale the local uncertainty distribution. Because it is an uncertainty distribution, the kriging variance should not be changed because it defines possible error range in the kriging estimate. Especially for the EDDSS, the kriging variance drop fast and go zero at the last cell. Changing the kriging variance would be wrong and resulting a very high or low variance in the simulation results.

Another possible reason is the kriging estimate. In the Gaussian space, if a nearby previously simulated value is high, the kriging estimate is likely high but not close due to the small data range in Gaussian space. In the direct space, if a nearby previously simulated value is high, the kriging estimate could be very high. Using the CCDF tables, it has more chance to draw a high value. And this high value will affect the subsequent simulation of cells. Besides, the screening effect of block datum in EDDSS could cause high or low kriging estimates. Thus, the kriging estimates are higher than the CCDF table means. And the DSS simulation realizations often show patching style images, with big changes from adjacent cells.

In DSS, we have to keep the drawing values close to the kriging estimate. Kriging variance indicates the error variance in the kriging estimate, and also the missing variance in kriging results. Keep the drawing value in the range of two standard deviations away from the kriging estimate maybe can add back the missing variance. However, to reproduce the histogram, we still need to know the distribution shape.

Lognormal distribution can be reproduced using analytical method (Manchuk, *et al.* 2005).

A Large Example of Exact Downscaling

The exact downscaling techniques can be implemented in four steps: 1) data assembly, 2) Variogram Modeling, 3) exact downscaling with kriging or DSS, 4) checking the results. Let is show them with a large example.

Data Assembly

Block data are obtained from Bayesian Updating results or a very reliable source. It should cover the entire interested area for fine scale model. Well (point) data are selected based on the availability. The block and well data should be checked to ensure their consistency.

In this example, the interested area is about 4 sections (each section is 1 square mile): 3200m x 3200m. The porosity block data is taken from a large 2-D prior model at the scale of 50m by 50m. The map and histogram of the block data are shown in Figure 6. There are 16 wells in the study area with log data. The locations of the 16 wells and the histogram of the well log data are shown in Figure 7. They are consistent with each other.

Variogram Modeling

Variogram need to be calculated from the point data. If there is no enough point data available, the variogram model used for Bayesian Updating 2-D mapping can be adopted. The horizontal and vertical variograms are calculated from the well log data, and the variogram models are shown in Figure 8.

Exact Downscaling

The 2-D block data is extended to a 3-D model using both EDDK and EDDSS methods. The block data are downscaled in the vertical direction to generate a 3-D model at scale of 50m by 50m by 1m. The 3-D porosity model from EDDK is shown in Figures 9 and 10. And the 3-D porosity model from EDDSS is shown in Figures 11 and 12. The map of block porosity (top left) and all odd number x-y view slices of the 3-D model are shown together (Figures 11 and 9). The dash lines in the block porosity map indicate the locations of the x-z and y-z cross sections. The x-z and y-z cross sections of the 3-D Model are shown in Figures 12 and 10. The dish lines in the x-z and y-z cross sections indicate the wells. The well data are reproduced. And the horizontal continuity is clearly shown in the cross sections. EDDSS realization has more variation than the EDDK result.

Checking Results

The 3-D model is converted back to 2-D model using the arithmetic averaging of each column. The results are plotted together with the original 2-D porosity map and the cross plot of the two grids in Figure 13. After downscaling to a 3-D model and upscaling back to a 2-D model, the new 2-D map is exactly same as the original map. The consistency between the two models indicates that the scaling method is exact, and provide more confident in dealing with multiscale data.

The histogram of the downscaled porosity is shown in Figure 14. It shows the mean is same as the mean of block data (Figure 6), and the maximum and minimum values are same as the point data (Figure 7). Because kriging estimates are smooth, a standard deviation of 0.0282 falls between the block data (0.0161) and the point data (0.0561).

Conclusion

The simple block cokriging and DSS using the block and point data can exactly reproduce the block data. Therefore, these methods can be used to downscale the block data. Some features and implementation aspects of the EDDK and EDDSS are given.

References

C.V. Deutsch and A.G. Journel, *GSLIB Geostatistical Software Library and User's Guide*, Second Edition, Oxford University Press, 1998

C.V. Deutsch, T. Tran, and Y. Xie, *A preliminary report on: An approach to ensure histogram reproduction in direct sequential simulation*, Technical report, Center for Computational Geostatistics, University of Alberta, Edmonton, AB, March 2001

A.G. Journel and Ch.J. Huijbregts, *Mining Geostatistics*, Academic Press Limited, 1978

B. Oz and C. V. Deutsch, *A Short Note on the Proportional Effect and Direct Sequential Simulation*, Centre for Computational Geostatistics, University of Alberta, Report 4: 2001/2002

B. Oz, C. Deutsch, T. Tran, and Y. Xie, *A Fortan 90 program for direct sequential simulation with histogram reproduction*, Computers & Geosciences, page submitted, 2001

M.J. Pyrcz and C.V. Deutsch, *Building Blocks for Direct Sequential Simulation on Unstructured Grids*, Center for Computational Geostatistics, University of Alberta, Report 4: 2001/2002

W. Xu and A.G. Journel, *DSSIM: A General Sequential Simulation Algorithm*, Stanford Center for Reservoir Forecasting, Stanford, CA, May 1994

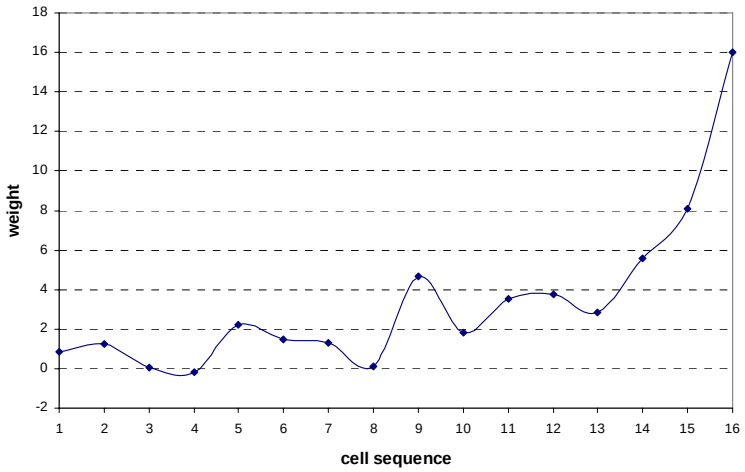


Figure 1: The kriging weight of block datum changes with simulation path.

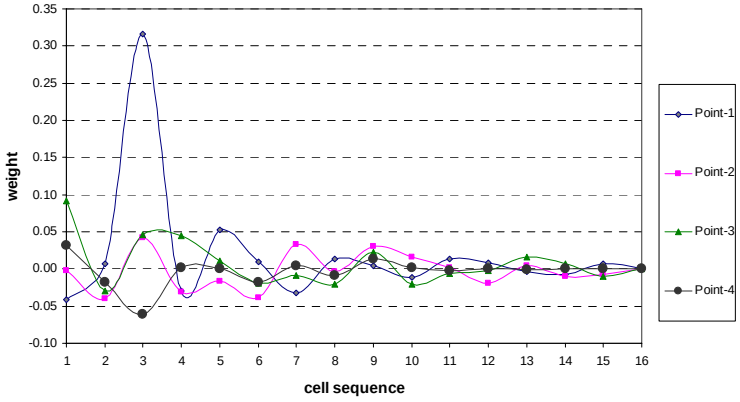


Figure 2: The Kriging weights of point data changes with simulation path.

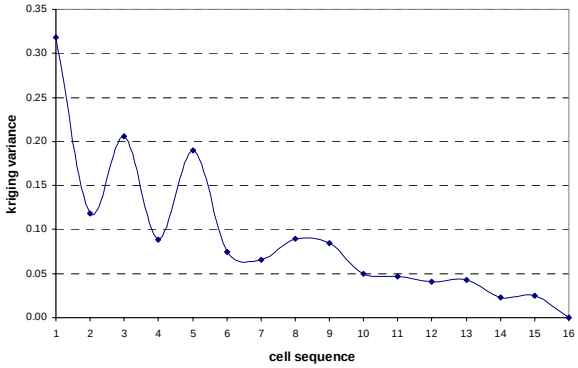


Figure 3: The kriging variance changes with simulation path.

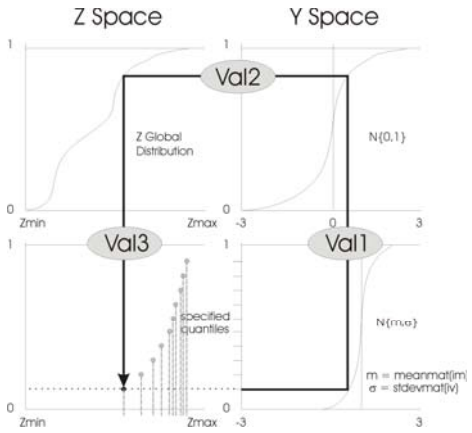


Figure 4: The graphical representation of the transformations applied to calculate the local distributions of uncertainty (Source: M.J. Pyrcz and C.V. Deutsch, 2001)

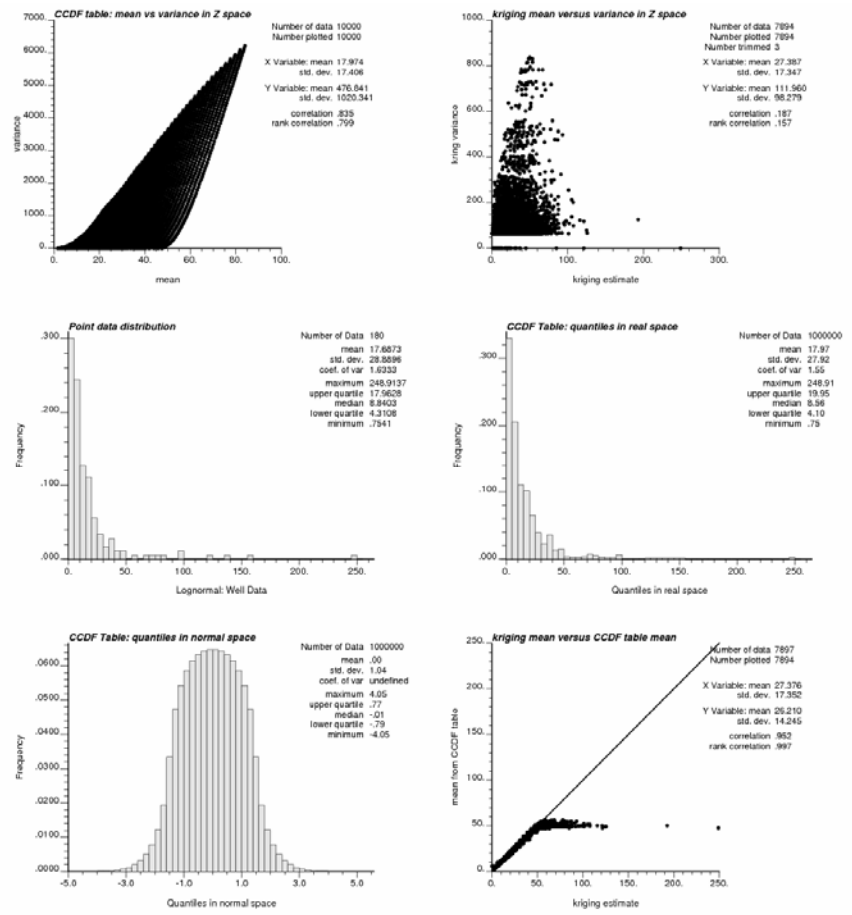


Figure 5: The CCDF lookup-table method results with a lognormal global distribution. The mean limit is 1.5 and variance limit is 1 to reach a standard normal distribution (bottom left).

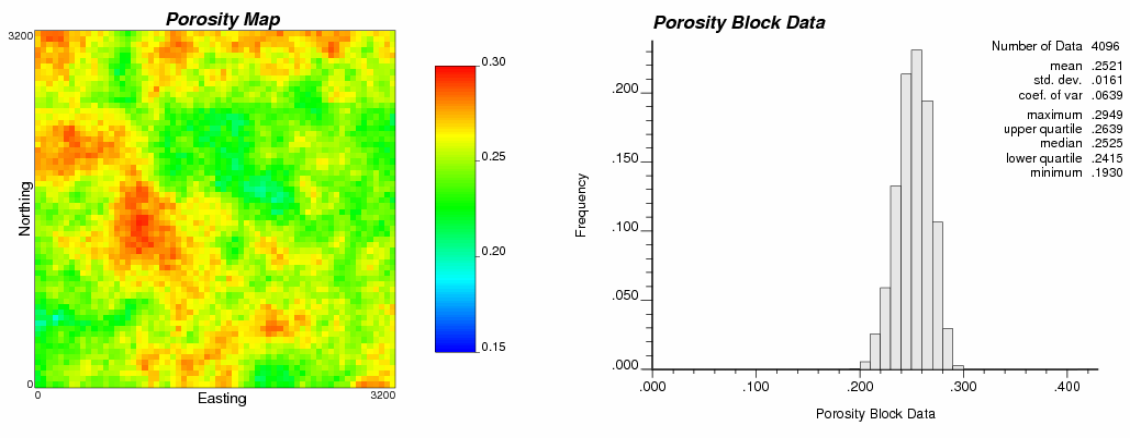


Figure 6: 2-D map of porosity block data at scale of 50m x 50m x 10 m.

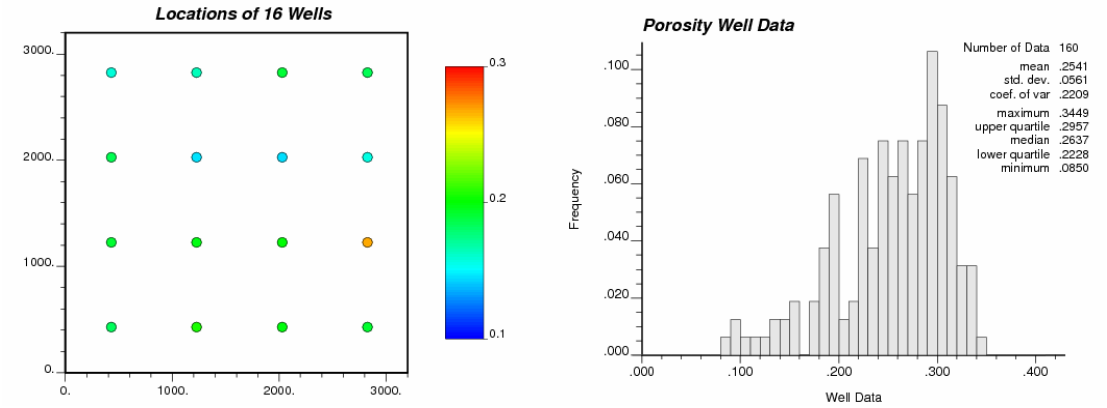


Figure 7: The location map of the 16 wells and the histogram of the well log data.

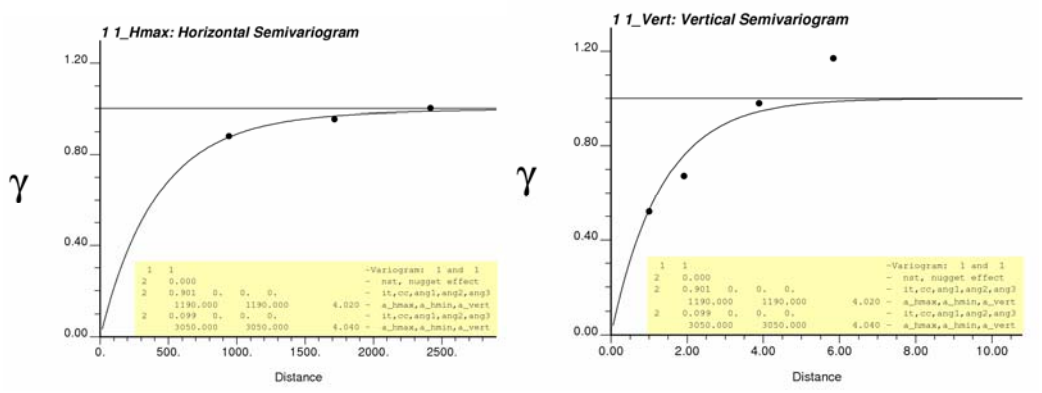


Figure 8: The horizontal and vertical variograms and variogram models.

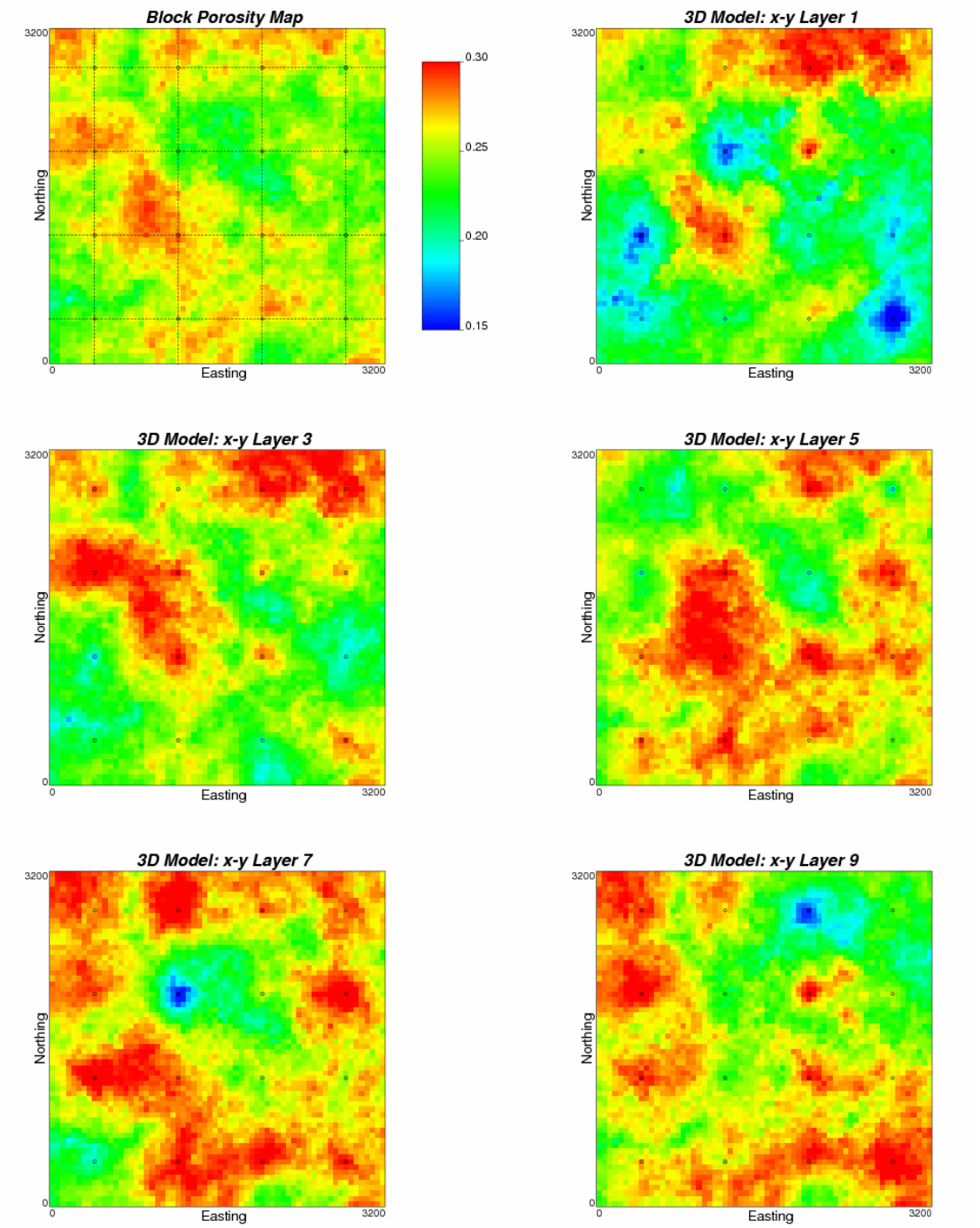


Figure 9: The 2-D porosity map and all odd number x-y view slices of the 3-D model from EDDK. The 3-D porosity is at scale of 50x50x1m. The dash lines in the porosity map (the top left) indicate the locations of the x-z and y-z cross sections.

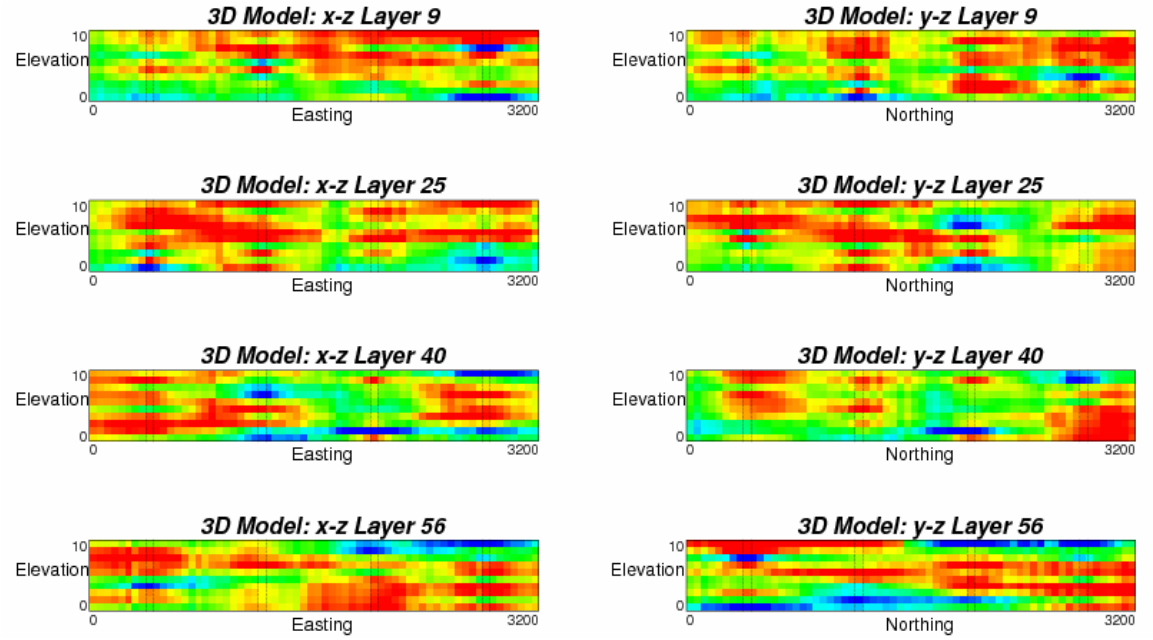


Figure 10: The x-z and y-z cross sections of the 3-D porosity model from EDDK. The 3-D porosity is at scale of 50x50x1m. The dish lines in the x-z and y-z cross sections indicate the wells. The horizontal continuity matches the well data.

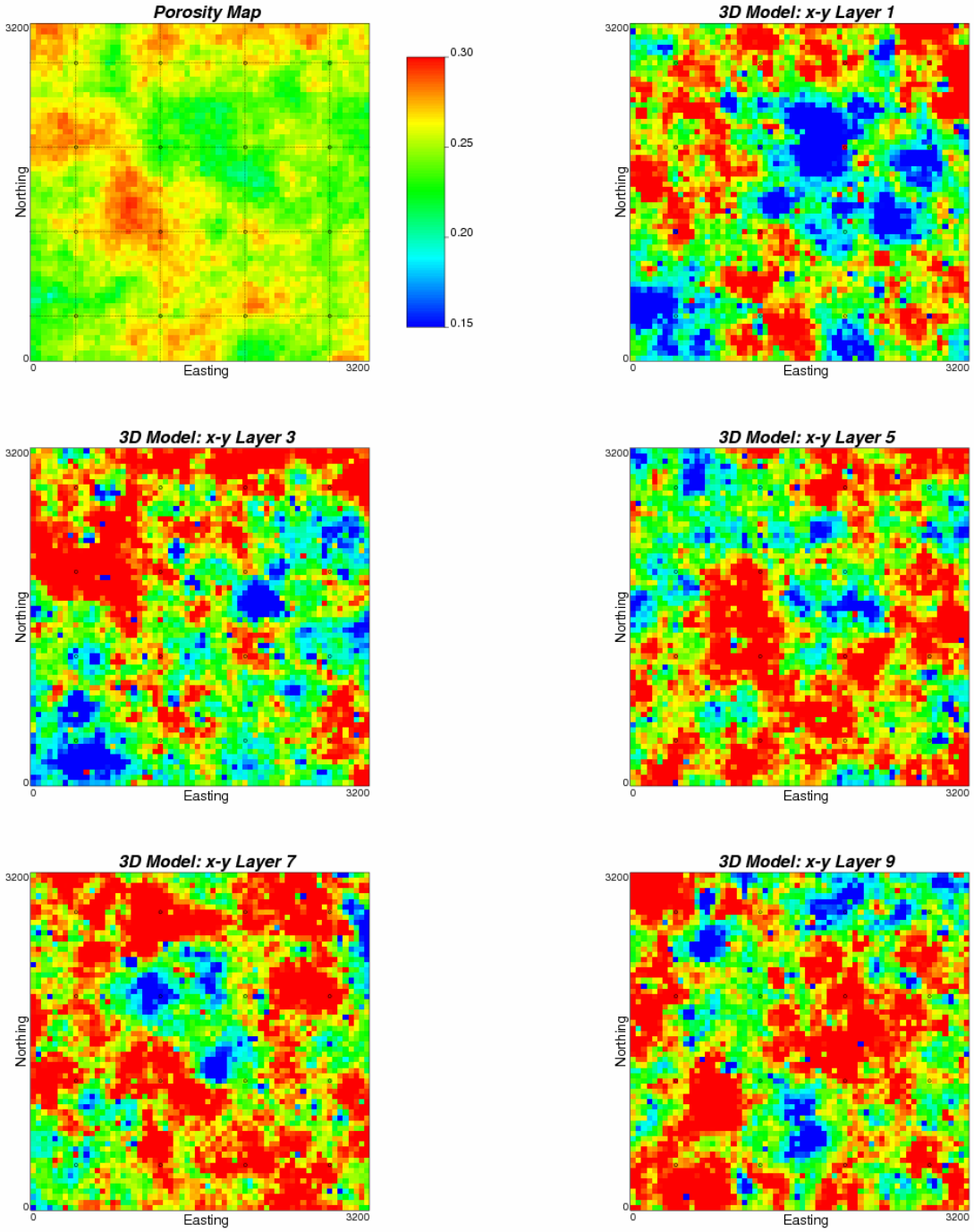


Figure 11: The 2-D porosity map and all odd number x-y view slices of the 3-D model from EDDSS. The 3-D porosity is at scale of 50x50x1m. The dash lines in the porosity map (the top left) indicate the locations of the x-z and y-z cross sections.

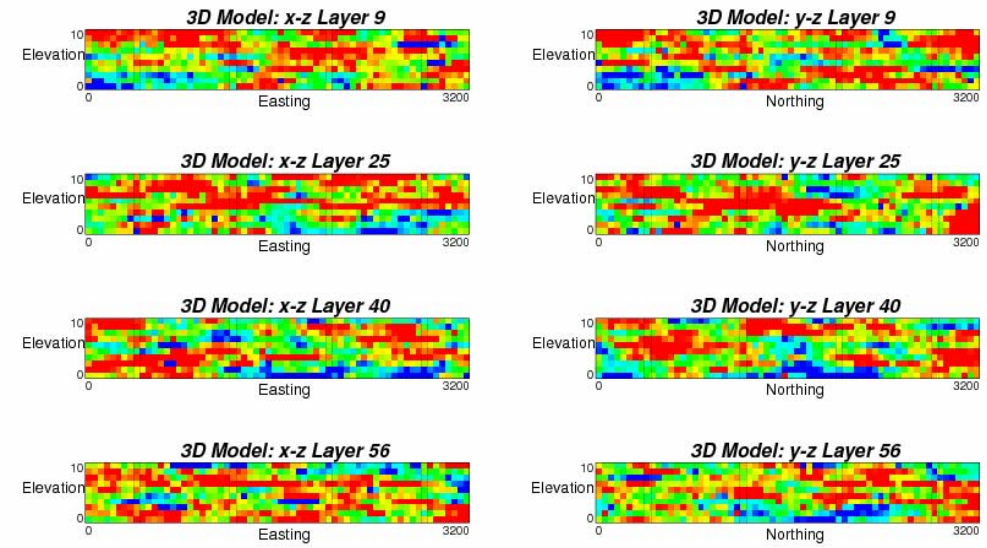


Figure 12: The x-z and y-z cross sections of the 3-D porosity model from EDDSS. The 3-D porosity is at scale of 50x50x1m. The dish lines in the x-z and y-z cross sections indicate the wells. The horizontal continuity matches the well data.

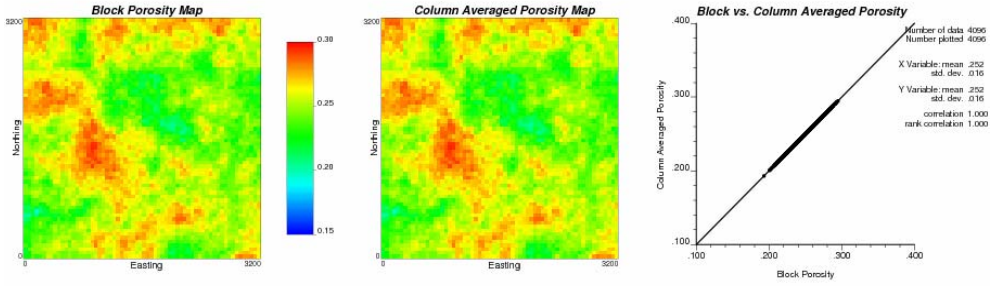


Figure 13: The block 2-D porosity map (left), the map of the column averaged porosity from the 3D model (middle), and the cross plot of the two grids.

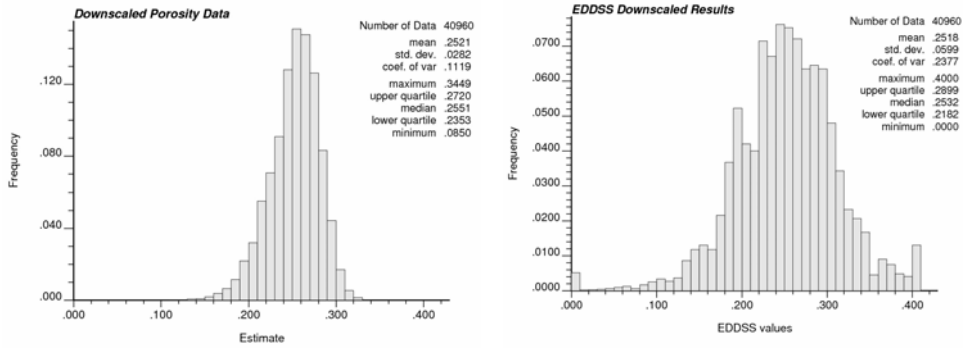


Figure 14: The histogram of fine scale porosity estimates from EDDK (left) and EDDSS (right).