Controlling the uncertainty in reservoir stochastic simulation

Cui Yong^{1*}, Chi Bo², Chen Guo³, Ouyang Cheng³ and Xia Bairu¹

¹ China University of Geosciences, Beijing 100083, China

² Research Institute of Daqing Oilfield Company Ltd., Heilongjiang 163712, China

³ Geological Exploration & Development Research Institute, PetroChina Chuanqing Drilling Engineering Company, Chengdu, Sichuan 610051, China

© China University of Petroleum (Beijing) and Springer-Verlag Berlin Heidelberg 2010

Abstract: Unexpected noise in reservoir stochastic simulation realization may be too high to make the realization useful, especially when there is a lack of hard data. Through discussing the uncertainties, we present two ways to control the uncertainty ratio that is brought by the algorithm of stochastic simulation. By reasonably reducing the random value of the stochastic simulation result, the unexpected values introduced by the residual that associates with random series can be controlled. Another way when the data disperse unevenly is to control the stochastic simulation order by grouping the points that need to be simulated to make those points which can be simulated by more neighborhood hard data calculated first. Both methods do not go against the core stochastic simulation algorithm.

Key words: Reservoir stochastic simulation, hard data, Kriging algorithm, residual, realization

1 Introduction

In the past several decades, two-point geostatistical methods have been developed intensively. Reports of these are dispersed in a variety of technical journals and books. At the same time, the methods were widely used in oil production industry and proved useful (Tolosana-Delgado et al, 2008; Hauge et al, 2007; Deutsch, 2002; Hohn, 1999).

Because of the characteristics of geology, there is always a lack of hard data in reservoir stochastic simulation. Many results are derived from the sparse data from drilling. Therefore, many researchers try hard to study the methods to analyze the accuracy of modeling results (Yamamoto, 2005; Pardo-Iguzquiza and Dowd, 2005; Patriarche et al, 2005; Rivoirard, 2004; Marchant and Lark, 2004; Yao, 2004; Leuangthong and Deutsch, 2004; Caers, 2000). Most of those works focused on properly fitting the variogram or eliminating data trends, and little attention was paid to the sparse data themselves. When the data are transformed by Q-Q plot (Deutsch, 2002) to normal distribution, the data are also declustered to have the most representative features (Deutsch, 2002). We also have something to consider, that is, the relationship between simulated points and hard data.

As the realizations of stochastic simulation have great influence on geological classification of the reservoir properties (Verly, 2005), whether one realization can be used or not has to be determined first. When useful information is

masked by random noise in one realization, the contribution for evaluating the result is limited or even negative. Something has to be done towards discounting the random residual part in the simulation (Hu, 2008).

2 Overview of geostatistical reservoir modeling

The key concept in geostatistical reservoir modeling is Kriging. That is the core of the two points geostatistics. By using variograms to present the spatial continuity of geological data, this method builds data pairs in a given direction. Through the relationship of the data pairs, we can estimate or simulate the data around or near the data pairs. These are the most simple and also most important steps of reservoir modeling.

The variogram is defined as:

$$2\gamma(\vec{h}) = E \left\{ \left[Y(\vec{u}) - Y(\vec{u} + \vec{h}) \right]^2 \right\} \quad (1)$$

where γ is half variogram, $Y(\vec{u})$ is the hard data value at the location $\vec{u}$, $\vec{h}$ is the distance from the location $\vec{u}$.

The covariance is defined as:

$$C(\vec{h}) = E \left\{ Y(\vec{u}) \cdot Y(\vec{u} + \vec{h}) \right\} - \left[E \left\{ Y(\vec{u}) \right\} \right]^2 \quad (2)$$

*Corresponding author. email: cui1019@sina.com, kaibplatform@163.com

Received August 31, 2009

The relationship between the variogram and the covariance is:

$$C(\vec{h}) = C(0) - \gamma(\vec{h}) \quad (3)$$

The linear estimator at a location is:

$$Y^*(\vec{u}) = \sum_{\alpha=1}^n \lambda_{\alpha} \cdot Y(\vec{u}_{\alpha}) \quad (4)$$

where $Y^*(\vec{u})$ is the estimated value, λ_{α} is the weight applied to the hard data value. The error variance for the estimator is:

$$E \left\{ \left[Y^*(\vec{u}) - Y(\vec{u}) \right]^2 \right\} = \sigma_E^2(\vec{u}) \quad (5)$$

By expanding this equation, we can find out that the error variance is:

$$E \left\{ \left[Y^*(\vec{u}) - Y(\vec{u}) \right]^2 \right\} = \sum_{\alpha=1}^n \lambda_{\alpha} \sum_{\beta=1}^n \lambda_{\beta} C(\vec{u}_{\beta} - \vec{u}_{\alpha}) - 2 \cdot \sum_{\alpha=1}^n \lambda_{\alpha} C(\vec{u} - \vec{u}_{\alpha}) + C(0) \quad (6)$$

The partial derivative of the other two parts of the equation with respect to weight λ_{α} is calculated as:

$$\frac{\partial \left[\sigma_E^2(\vec{u}) \right]}{\partial \lambda_{\alpha}} = 2 \cdot \sum_{\beta=1}^n \lambda_{\beta} C(\vec{u}_{\beta} - \vec{u}_{\alpha}) - 2 \cdot C(\vec{u} - \vec{u}_{\alpha}) \quad (7)$$

$$\alpha = 1, 2, \dots, n$$

By minimizing the estimation variance these may be set to zero:

$$\sum_{\beta=1}^n \lambda_{\beta} C(\vec{u}_{\beta} - \vec{u}_{\alpha}) = C(\vec{u} - \vec{u}_{\alpha}) \quad \alpha = 1, 2, \dots, n \quad (8)$$

The system equations with n unknown weights are known as the simple Kriging (SK) system.

3 Uncertainty factors in reservoir stochastic simulation

3.1 Kriging variance

During stochastic simulation, the Kriging variance is defined as:

$$\sigma_{SK}^2 = C(0) - \sum_{\alpha=1}^n \lambda_{\alpha} C(\vec{u}, \vec{u}_{\alpha}) \quad (9)$$

Ideally, this variance is nearly zero. However, this will never happen in geology reservoir stochastic simulation. On the other hand, we have to add the variance to the result by randomly drawing a value from a normal distribution with mean of 0 and variance of σ_{SK}^2 .

$$Y_s(\vec{u}) = Y^*(\vec{u}) + R(\vec{u}) \quad (10)$$

When σ_{SK}^2 is high, the residual $R(\vec{u})$ is high too, and because the value is drawn randomly, the uncertainty is

important at the location where the Kriging variance is high.

3.2 Stochastic simulation order

When the hard data are located evenly, it is unnecessary to consider which point should be simulated first and which one should be simulated later. However, in geology, this is not a general phenomenon, and geological data are almost always dispersed around the high point of structures or those sites with better exploration or development prospects.

In simulation, generally, the more hard data available that affect the simulated point, the more reliable the simulation result we can get. So, it is meaningful to control the random simulation order reasonably. It is good to simulate those points that have more surrounding hard data first. Practically, it is recommended to separate the points which need to be simulated into several groups by the amount of hard data which can be searched. Points in the group with bigger searched number should be simulated first. In practice, this is a balance of Kriging algorithm and reliability of the stochastic simulation result.

4 Methods

4.1 Controlling Kriging variance

Properly reducing the random part of the simulation result is a way to control the uncertainty. As defined in the Kriging algorithm, the residual is drawn completely randomly. Sometimes this part of the result may exceed the Kriging value. If this happens, the result map looks more like random noise than useful information.

For conditionally reducing the residual, we process the results with residuals exceeding the Kriging estimated values. When the simulated result is composed of more than half non-random part, the geological trend of the data will be retained in the realization. The residual of a simulated point is controlled by σ_{SK}^2 , which has a direct relationship with Kriging weights (Eq. (9)). As discussed above, if the Kriging variance is high enough, the random part, that is, the residual, will be high enough to make the simulation result useless.

Normally but not absolutely, if there are more hard data weights for one simulated point, the Kriging variance tends to be higher. So, when the residual is higher than half the Kriging estimated value, we divide the Kriging variance by the number of the linear equations, then again, draw a value from a normal distribution with mean of 0 and variance of σ'^2 randomly, where σ' is defined as:

$$\sigma'^2 = \sigma_{SK}^2 / n \quad (11)$$

where n is the number of the Kriging weights. In this way, in most cases the random part drawn from the normal distribution will be less than the Kriging estimated value, and the geological trend will not be masked by the random part in the realization.

4.2 Controlling simulation order

The other way to get the same result is to use the hard data in a more effective way. When simulating a point, normally,

the more hard data that are searched, the less possible it is to get an unexpected result. Some researchers noticed that the order of simulation affects the result. Liu and Journel (2004) used conditional probability of grids data from training image to reduce the randomness in multiple-point simulation. In this paper, we use the amount of the hard data searched during simulation to fulfill the same job.

By using hard data more effectively, we have to do the search work twice. The first search is to find out the amount of hard data which could be used when simulating a point later, and the second time is for simulating the point.

The steps of the method are:

- Mark the points to be simulated;
- Get the amount of hard data of each point that could be used when it is simulated;
- Sort the points by the amount of hard data from step b, higher first;
- Divide the sorted points to several groups serially (the number of groups can be changed, normally, there should have enough data in each group for random simulation);
- From first to last, for each group, randomly select a point to simulate.

5 Examples

5.1 Hard data

We take one set of data about sandstone layer thickness from an oilfield as an example. Fourteen data points disperse unevenly in the 36 grids and 23 grids do not have any data. The grid size is 155m×155m (Fig. 1).

Fig. 1 Location of sandstone layer thickness data

The contour map (For comparing the original data and the simulation results, all maps in this paper are drawn with Surfer 8.0, Golden Software Inc, Colorado.) of the original data is shown in Fig. 2. The map algorithm used here is inverse distance. Because of sparse original data, the map looks very smooth. However, this map has little geological meaning, and it shows the data only.

Two parametric variograms are used to describe continuities of two directions. The main direction is 30° clockwise from north, and its variogram is represented by the red line in Fig. 3. The other blue line describes the variogram of 120° clockwise from north. The parameters of parametric variograms are shown in Table 1. They are set by considering the facies analysis and adjacent field analogy. All the simulation presented in this paper use the two variograms.

Fig. 2 Contour map drawn by the original hard data only (Data location is shown in Fig. 1)

Fig. 3 Variograms used in the example

Table 1 Parameters corresponding to variograms shown in Fig. 3

Variogram	Type	Horizontal range	Nugget effect	Sill
NE 30°	Gaussian	360 m	--	0.9
NE 120°	Spherical	300 m	0.05	0.9

5.2 Uncertainty of SGS stochastic simulation

Without any additional processing of the hard data, we use sequential Gaussian simulation (SGS) to simulate the original hard data. The result is shown in Fig. 4. To compare with the estimation result that reflected the geological trend of the data described by the variograms (Fig. 5), there are too many unexpected simulation results, and the map looks very noisy because of the compensation of the loss variance.

For simulation, one thing has to be considered, that is whether it is proper to treat all data the same way. Obviously, this realization is too noisy to be used in geology analysis, especially the simulated data located on top left area where

Fig. 4 One SGS realization of data shown in Fig. 1 (The darker color represents smaller value, and the brighter color represents higher value)

Fig. 5 Kriging estimation of the data shown in Fig. 1

the density of hard data is low. Then, we propose our methods to solve the problem. Consider the stochastic simulation algorithm described above, if we control the random part of the result, that is simulation residual, we can manage the uncertainty in the result.

5.3 Reducing the residual ratio

Normally, during simulation, the amount of hard data that could be used to calculate the simulation result may vary from one point to another. And again, the more hard data found to calculate the point, the more reliable the result we get. So we can reduce the simulation residual by this number. This number is easy to get, it is just the number of the Kriging weights. Then we use Eq. (11) to control the result. This time the map of the simulation result is shown in Fig. 6. The random series used is the same as that used in Fig. 4.

As shown in Fig. 6, the map shows much more geological meaning implied by the variograms in Fig. 5. In this case, we properly reduced the uncertainty part of the stochastic simulation result, the simulated points affected by few hard data added more residual.

Fig. 6 SGS realization by using neighborhood hard data amount balance

5.4 Ranking simulation points

Another way to reduce the uncertainty is the simulation order. We can separate the points to be simulated into several groups by the amount of hard data that can be searched in simulation. And then, from big number group to small one, the point in each group is simulated. The number of the group is selected according to the hard data amount. There is no theoretical way to appoint the group number. The number of the groups should be determined by user according to the data set. Fig. 7 shows one SGS realization of data pointed in Fig. 1. In this case, we set all simulated points into three groups first, and simulate points in each group.

Fig. 7 SGS realization by ranking simulated points into three groups by the hard data that the points can search

Fig. 7 shows more continuity than Fig. 4 but less than Fig. 6, because the hard data amount is low.

5.5 Combination of both methods

We can use the two methods together to control the uncertainty ratio of simulation. Fig. 8 shows one SGS realization that used three point groups and reduced the random residual by the equation number. The result indicates that we can manage the results of simulation more flexibly by the methods mentioned above.

Fig. 8 SGS realization by ranking simulated points into three groups and the stochastic simulation residual results were balanced by the Kriging weights number

6 Discussion

Sparsity of hard data is one critical problem in reservoir stochastic simulation. Sometimes we have to get the best result possible from the limited amount of hard data. The traditional methods of stochastic simulation can have a big uncertainty ratio in the result because the variograms used may vary from the practical one. The error will always exist if there are unknown data.

But we can manage the uncertainty ratio by geology concepts and the mathematical algorithm. This paper shows two ways to solve the problem. The first is reducing the uncertainty residual part by dividing the loss variance by the amount of the hard data the simulated points used. There is an implied condition which is the more searched data there exist the lower the loss variance. As we all know, this is normally true, but not absolutely correct. However, by considering the random regulation of the simulation, this will bring little error. Presently there is no critical way to analyze how the method could smooth the realization.

7 Conclusions

In this paper, we apply the amount of hard data searched

when simulating a point to reduce the uncertainty residual part properly. We propose two ways to reduce the uncertainty ratio during simulation. The result of improved SGS simulation indicates that each way affects the simulated value. By using the two methods we can get more reliable realization than the traditional SGS methods that could result in a noisy map.

References

- Caers J. Adding local accuracy to direct sequential simulation. *Mathematical Geology*. 2000. 32(7): 815-850
- Deutsch C V. *Geostatistical Reservoir Modeling*. Oxford: Oxford University Press. 2002. 101-241
- Hauge R, Holden L and Syversveen A R. Well conditioning in object models. *Mathematical Geology*. 2007. 39(4): 383-398
- Hohn M E. *Geostatistics and Petroleum Geology* (Second Edition). Netherlands: Kluwer Academic Publishers. 1999
- Hu L Y. Extended probability perturbation method for calibrating stochastic reservoir models. *Mathematical Geosciences*. 2008. 40(8): 875-885
- Leuangthong O and Deutsch C V. Transformation of residuals to avoid artifacts in geostatistical modelling with a trend. *Mathematical Geology*. 2004. 36(3): 287-305
- Liu Y H and Journel A. Improving sequential simulation with a structured path guided by information content. *Mathematical Geology*. 2004. 36(8): 945-964
- Marchant B P and Lark R M. Estimating variogram uncertainty. *Mathematical Geology*. 2004. 36(8): 867-898
- Pardo-Iguzquiza E and Dowd P A. Empirical maximum likelihood Kriging: the general case. *Mathematical Geology*. 2005. 37(5): 477-492
- Patriarche D, Castro M C and Goovaerts P. Estimating regional hydraulic conductivity fields—a comparative study of geostatistical methods. *Mathematical Geology*. 2005. 37(6): 587-613
- Rivoirard J. On some simplifications of cokriging neighborhood. *Mathematical Geology*. 2004. 36(8): 899-915
- Tolosana-Delgado R, Pawlowsky-Glahn V and Egozcue J J. Indicator Kriging without order relation violations. *Mathematical Geosciences*. 2008. 40(3): 327-347
- Verly G. Grade control classification of ore and waste: A critical review of estimation and simulation based procedures. *Mathematical Geology*. 2005. 37(5): 451-475
- Yamamoto J K. Correcting the smoothing effect of ordinary Kriging estimates. *Mathematical Geology*. 2005. 37(1): 69-94
- Yao T T. Reproduction of the mean, variance, and variogram model in spectral simulation. *Mathematical Geology*. 2004. 36(4): 487-506

(Edited by Hao Jie)