Paper 106, CCG Annual Report 11, 2009 (© 2009)

Multivariate Probability Estimation for Categorical Variables from Marginal Distribution Constraints

Yupeng Li and Clayton V. Deutsch

In geostatistics, the simulation of categorical variables such as facies or rock types requires a conditional distribution based on the information from nearby data locations. The conditional probability can easily be calculated from the multivariate distribution of the unsampled value and all data values. In this paper, a new multivariate probability estimation algorithm is proposed. The multivariate probability is estimated from lower order marginal probabilities that are known from the available data. The required conditional probability is then obtained directly by Bayes law. In this algorithm, the bivariate marginal probabilities are imposed to an initial multivariate probability as the constraints and satisfied by iteratively modifying the initial multivariate probability. The bivariate marginal probabilities are inferred from sampled locations, a training image or profiles along drill holes or wells. A sparse matrix is used to calculate the marginal probabilities from the multivariate probability, which significantly saves CPU time and makes this estimation algorithm practical. This algorithm can be extended to higher order (m>2) marginal probability constraints. The theoretical framework is developed and illustrated with a number of realistic examples.

Introduction

In reservoir management, a numerical reservoir model is always required for resource evaluation and reservoir flow simulation input for production forecasting and recovery calculations. Many factors will have an effect on numerical reservoir model. Most geostatistical practitioners agree that facies are one of the most important reservoir heterogeneities. In geology, a facies is a distinctive rock unit that forms under certain conditions of sedimentation, reflecting a particular process or environment (e.g. river channels, delta systems, submarine fans, reefs). Usually, facies are rock units that are somehow statistically homogeneous.

Although, the facies for each unsampled location is unique, it is viewed as a random variable that could take a set of possible values or states. The random variables are location dependent and donated with the location vector $\mathbf{u}$. The possible outcome of the random variable for the location is denoted as k and $k \in \{1, 2, ..., K\}$ which are mutually exclusive and exhaustive. Practically, the total number of facies K is less than 5. We can think of $\mathbf{u}=k$ as an event and $P(\mathbf{u}=k)$ is the probability of this particular event being true.

Multivariate probability: A specific spatial arrangement of n different locations will define a set of random variables $\{\mathbf{u}_i, i=1,...,n\}$. A state of the facies combination outcomes at each location for a set of n locations is denoted as $\{\mathbf{u}_i = k_i; i=1,...,n\}$. The uncertainty (our lack of knowledge) of whether or not a state $k_i \in \{1,2,...K\}$ exists at location $\mathbf{u}_i$ is characterized by a set of discrete multivariate probabilities $P(\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n; k_1, k_2, ..., k_n)$ written as:

$$\begin{aligned} P(\mathbf{u}_1,...,\mathbf{u}_n;k_1,...,k_n) &= \text{Prob}\{\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2, ..., \mathbf{u}_n = k_n\} \\ &= \text{Prob}\{\mathbf{u}_i = k_i, i=1,...,n\} \\ & \qquad k_i \in \{1,2,...K\};\ i=1,..,n \end{aligned} \tag{1}$$

The multivariate probability has the following properties:

$$0 \le P(\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n; k_1, k_2, ..., k_n) \le 1$$

$$\sum P(\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n; k_1, k_2, ..., k_n) = 1$$

For convenience, the notation for the multivariate probability will be simplified to $P(\mathbf{u}_i,, i=1,...,n; k_i, i=1,...,n)$. One characteristic of this discrete multivariate probability is its huge state space. The number of all possible states will be K^n , and these K^n states have different probabilities to exist. Consider there are 3 facies in the domain, and 20 sampled locations, then the multivariate probability state space will be $3^{20} = 3,486,784,401$.

Multivariate probability marginalization: Given the n-variate multivariate probability $P(\mathbf{u}_\ell = k_\ell, \ell = 1,...,n)$, different orders of marginal probability can be calculated based on the law of total probability. First order marginal probability $P_1(\mathbf{u}_i; k_i)$ as a univariate probability of each facies k_i for the location $\mathbf{u}_i$ is defined as:

$$P_1(\mathbf{u}_i; k_i) = \sum_{\substack{All\ \ell \\ with\ \mathbf{u}_i = k_i}} P(\mathbf{u}_\ell, \ell = 1,...,n; k_\ell, \ell = 1,...,n) \tag{2}$$

The univariate marginal probability satisfy $0 \le P_1(\mathbf{u}_i; k_i) \le 1$ and $\sum_{k_i=1}^{K} P_1(\mathbf{u}_i; k_i) = 1$.

Second order marginal or bivariate marginal probability $P_2(\mathbf{u}_i, \mathbf{u}_j; k_i, k_j)$ can be calculated from the multivariate probability as:

$$P_2(\mathbf{u}_i, \mathbf{u}_j; k_i, k_j) = \sum_{\substack{All\ \ell \\ with\ \mathbf{u}_i = k_i\ \&\ \mathbf{u}_j = k_j;\ i \ne j}} P(\mathbf{u}_\ell, \ell = 1,...,n; k_\ell, \ell = 1,...,n\ ) \tag{3}$$

The bivariate marginal probability satisfy $0 \le P_2(\mathbf{u}_i, \mathbf{u}_j; k_i, k_j) \le 1$ and $\sum_{k_i=1}^{K} \sum_{k_j=1}^{K} P_2(\mathbf{u}_i, \mathbf{u}_j; k_i, k_j) = 1$.

Following this logic, any *m* order marginal probability of the n-variates multivariate probability $P(\mathbf{u}_1,...,\mathbf{u}_m; k_1,...,k_m)$ simplified as $P_m(k_1,...,k_m)$ can be calculated as:

$$P_m(\mathbf{u}_1,...,\mathbf{u}_m; k_1,...,k_m) = \sum_{\substack{All\ \ell \\ with\ \mathbf{u}_1 = k_1,...,\mathbf{u}_m = k_m; m \le n}} P(\mathbf{u}_\ell, \ell = 1,...,n; k_\ell, \ell = 1,...,n\ ) \tag{4}$$

In expression(4), the m-variate marginal probability can be any *m* arbitrary subgroup of the *n* random variables.

Conditional probability: we are often interested in estimating the probability that facies k_i prevails at location $\mathbf{u}_i$ based on the neighboring data. The unsampled location is denoted as $\mathbf{u}_1$ (could be any location in the domain, just for index convenience) in these n neighboring locations $\{\mathbf{u}_i = k_i, i = 1,...,n\}$. The probability of facies k_1 for unsampled location $\mathbf{u}_1$ will be: $P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)$ and can be calculated from the multivariate probability based on Bayes law:

$$P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n) = \frac{P(\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)}{P(\mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)} \tag{5}$$

In this equation, the numerator is an n-variates multivariate probability, while the denominator is the (n-1)-variates marginal probability of this n-variates multivariate probability.

Using the multi-variates marginal probability calculation as expressed in(4), the conditional probability in (5) will be calculated as:

$$P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n) = \frac{P(\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)}{P(\mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)} = \frac{P(\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)}{\sum_{k_1=1}^{k_1=K} P(\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)} \quad (6)$$

As stated in Equation(6), the inference of this conditional probability distribution is the central problem of geostatistics. There are several established and new approaches to infer the conditional probability directly or indirectly, such as indicator kriging or multiple point geostatistics. In indicator approach (Goovaerts, 1994; Journel, 1983), the conditional probability multivariate probability is calculated as a linear combination of indicator data:

$$P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n) = \sum_{i=2}^{n} \lambda_i P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_i = k_i) \quad (7)$$

The weights characterize the dependence between the sampled locations $\mathbf{u}_i$ and the information from sampled locations to unsampled location.

In the multipoint statistics approach, the required n-variate multivariate probability is constructed from a training image. The training image represents the heterogeneity characteristics that the geologist expects to see in the study area. Then, the required conditional probability in expression (6) can be calculated by just counting the relatively states number occurred in the training image:

$$P(\mathbf{u}_1 = k_1 \mid \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n\} = \frac{Count(\mathbf{u}_1 = k_1, \mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)}{Count(\mathbf{u}_2 = k_2,...,\mathbf{u}_n = k_n)} \quad (8)$$

Those two approaches are reasonable, but the linear probability combination approach of in indicator approach may not result in models that show the appropriate level of spatial detail. The stationary assumption in multipoint geostatistics may cause some difficulty. Furthermore, because of the huge state space, it is difficult to get enough replication for higher order multivariate probability from a single training image. Directly estimating the multivariate probability from lower marginal probabilities is proposed here to infer the multivariate probability in expression (6) directly. The lower marginal probabilities will have minimal stationary requirement and the nonlinear combination will be more appropriate in many cases.

Multivariate probability estimation from marginal constraints

When the random variables $\{\mathbf{u}_i = k_i; i = 1,...,n\}$ are independent, the multivariate probability is given by:

$$P(\mathbf{u}_1,...,\mathbf{u}_n; k_1,...,k_n) = \prod_{i=1}^{n} p(\mathbf{u}_i; k_i),\ k_i = 1,...,K \quad (9)$$

In most spatial statistics case, the random variables are dependent. The data information redundancy should be considered in the estimation process. Lower marginal probabilities are used as a tool to characterize the information and the redundancy between the sampled locations and the unsampled location. In this approach, the lower marginal probability is used as constraints to iteratively modify the estimated multivariate probability until they are all satisfied. The direct multivariate probability estimation algorithm (DMPE) illustrated in pseudo-program language in general case is:

Begin

Input: *the univariate probability* $\mathbf{p_1}$ *and the known m order marginal probability* $\mathbf{Pm}^{\text{Target}}$

Generate an initial multivariate probability $\mathbf{P}^*$ *with the independent assumption:* $\mathbf{P}^* = \prod \mathbf{p_1}$

Repeat

- *Use the* $\mathbf{P}^*$ *as the initial multivariate probability estimation* $\mathbf{P}^{\delta}$*;*
- *calculate the current marginal* $\mathbf{Pm}^{\text{Calc}}$ *from the current estimated multivariate probability* $\mathbf{P}^{\delta}$*;*
- *calculate a modify factor vector* $\mathbf{F} = \frac{\mathbf{Pm}^{\text{Target}}}{\mathbf{Pm}^{\text{Calc}}}$ *with the known marginal probability and the current estimated marginal probability;*
- *calculate a new updated multivariate probability with the modify factor:* $\mathbf{p}^{\delta+1} = (\mathbf{F} \times \mathbf{p}^{\delta}) \times C$
- *give* $\mathbf{P}^{\delta+1}$ *to* $\mathbf{P}^{\delta}$ *for next iteration;*

Until: *the change of the multivariate probability is stable:* $\Delta O \approx \varepsilon$ With $O = \| \mathbf{p}^{\delta+1} - \mathbf{p}^{\delta} \|$

Output: *the final estimated multivariate probability* $\mathbf{P}^{est}$

End

In the modification process, C is used to do normalization to make $\mathbf{P}^{\delta+1}$ a licit probability.

Small Numerical Example

To illustrate the DMPE process, consider the following spatial simulation problem. One a 1×3 grid, three categories may exist at any one of them. In this small case $(n = 3, K = 3)$, the total states number will be $3^3 = 27$, all the possible states are shown in Figure 1. Each of the possible states will have a specific probability to exist which is characterized by a multivariate probability. In order to store and index all the probabilities more efficiently, the multivariate probabilities are ordered in a one dimension array and given a unique index to each of them. The K^n states indexed from 1 to K^n will be calculated using the following index function:

$$f(k_\ell, \ell = 1, \ldots, n) = 1 + \sum_{\ell=1}^{n} (k_\ell - 1) \times K^{\ell-1} \tag{10}$$

In Equation(10), $k_\ell = 1, 2, \ldots, K$. It is obtained by ordering and coding all the categories into an integer set according to the order, but how to order the categories does not matter. Table 1 is the index calculation example for this case.

In this small example, the total number of second order combinations in three grids will be ${}_3C_2 = 3$, and each of them may have 9 outcomes, the outcomes of grid 1 and 3 are shown in Figure 2. The similar outcomes for combination of grid 1 and 2, or grid 2 and 3, so totally there will be 27 outcomes. Generally, for bivariate marginal probability, the total state number will be $K^2 \times {}_nC_2$. Denote the multivariate probability is $\mathbf{P} = [p_1, p_2 \ldots, p_{27}]^T$ which the probability for each state to exist and also denote the bivariate probability as $\mathbf{P2} = [b_1, b_2, \ldots, b_{27}]^T$. From equation(3), in this case, each bivariate probability can be calculated from the multivariate probability as:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ b_{27} \end{bmatrix} = \begin{bmatrix}
1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0 \\
0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0 \\
0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1 \\
1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0 \\
0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0 \\
0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1 \\
1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1
\end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ p_{27} \end{bmatrix}$$

As showing in above matrix equation, each bivariate probability assignment is the sum of the subset coming from the full table of the multivariate probability. Their relationship is characterized by the indicator index matrix $\mathbf{I2}$, and it can be written in matrix form as:

$$\mathbf{P2} = \mathbf{I2} \times \mathbf{P} \tag{11}$$

As the number of locations or grid nodes increase, the dimension of matrix $\mathbf{I2}$ increases dramatically. If there are 10 locations and 3 facies, the dimension of matrix $\mathbf{I2}$ will be 405×59049. And the non-zero elements number in each row is $K^{n-2} = 3^{10-2} = 6561$ which compose a sparse matrix. The column numbers for the nonzero value 1 in the sparse matrix depend on which pair of bivariate marginal is calculated. The marginalization computation will be a linear sparse matrix computation which is very fast.
If the bivariate probability for each two locations was given, it will be shown next that the multivariate probability can be estimated. Note, in the above (DMPE) algorithm, the lower order marginal probability can be any higher order than bivariate marginal probability. In general, denote $\mathbf{Pm}$ as *m* order marginal probability vector, it can be calculated from the multivariate probability as:

$$\mathbf{Pm} = \mathbf{Im} \times \mathbf{P} \tag{12}$$

The matrix $\mathbf{Im}$ is the indicator index matrix, the dimension is $K^m \binom{n}{m} \times K^n$ with $\binom{n}{m} = \frac{n!}{m!(n-m)!}$.

Assume the multivariate probability for the state outcome is already known as in Table 2. From the known multivariate probability, the bivariate probability can also be calculated and listed in Table 2. The input for DMPE algorithm is the univariate probability and bivariate probability. The estimation result is also listed Table 2 which is shown that the known multivariate probabilities are reproduced fairly well.

The DMPE algorithm conceived in the effort of geostatistics facies modeling algorithm improvement is coincided with one multivariate probability approximation in information theory research proposed by Ku and Kullback (1969). In their approach, based on the maximum entropy theory, the multivariate probability is obtained from the lower marginal probability by a straightforward iterative algorithm which is very similar to the proposed DMPE algorithm in this paper. In Ku and Kullback approach, for a multivariate probability with N variables K discrete categories, the pair wise constraints will require $K^N/K^2=K^{N-2}$ terms from the full multivariate terms. This operation will be very CPU intensive for large N (dependence on K). Many authors recognized this approach as a fantastic approach but noted that the computational complex is a major barrier(Freeman, 1971; Saerens and Fouss, 2004) and thus few practical applications.

While in DMPE approach, each single state of the full states set can be easily indexed. Also, using the index indicator transform, the marginalization computation can be done in a very fast linear operation style with a small storage requirement by taking advantage of the sparse matrix. This higher efficient index and sparse matrix computation in the iterative scheme successfully make this powerful estimation more applicable in practical problems. Table 3 is the comparison between these three methods using the same data as those used by Ku and Kullback (1969).

Reservoir Example

The above DMPE algorithm can be used in the traditional cell-based sequential simulation approach (Deutsch and Journel, 1998) where all simulation grid cells are visited only once along a random path and simulated cell values become conditioning data for cells visited later in the sequence.

Any unsampled grid cell $\mathbf{u}$ visited along the random path is simulated as follows:

1. Look for the n conditioning data (original well data or previously simulated cell values) closest to $\mathbf{u}$;
2. Based on the distance between every two locations, retrieve bivariate probability from the experimental bivariate probability diagram, see the coming explanation;
3. Using the DMPE algorithm to estimate the multivariate probability;
3. The estimated conditional probability of each facies at $\mathbf{u}$ is computed using the Bayes law based on the facies at each conditioning data location, see equation(6);

4. Draw a simulated facies value from the resulting local probability distribution using Monte-Carlo sampling, and assign that value to the grid cell $\mathbf{u}$ and go to the next unsampled location until all the unsampled location are visited.

The needed marginal probability in this DMPE approach can be obtained from a training image or geological analysis. The most easily obtained marginal is the univariate and bivariate marginal probabilities. The univariate marginal is the proportion of all the facies at each unsampled location. They could be the global proportion or locally varying proportions. In DMPE algorithm, the univariate probability will be used to calculate the initial estimation. The bivariate marginal probability for two locations is defined as $P_2(\mathbf{u}_i, \mathbf{u}_j; k_i, k_j)$ which can be obtained from the training image, well vertical profiles or from outcrops. From the training image in Figure 3, without thinking the different direction, as location $(\mathbf{u}_i, \mathbf{u}_j)$ apart from each other to a further distance, the isotropic bivariate probability between facies 1, 2 and 3 will compose a diagram as shown in Figure 4.
The bivariate probability diagram is close to the transition probability diagram (Li, 2007) that used in the new developed markov transition probability based geostatistics (Carle, 2000; Carle and Fogg, 1996; Li and Zhang, 2006).

Conclusion

Discrete multivariate probabilities can be directly estimated from the lower order marginal probability distribution constraints. The estimated multivariate probability works well with reasonable CPU time which promises practical use in the future. The use of m-variate (m>2) marginal constraints can be enforced in the iteration process given higher order marginal probabilities can be obtained. Practical 3D simulation and integrating sedimentary patterns from sequence stratigraphy will be the subject of future research.

References

CARLE, S.F., 2000. Use of a Transition Probability/Markov Approach to Improve Geostatistical of Facies Architecture. In., 250 Kilobytes pages ; PDFN.

CARLE, S.F. and FOGG, G.E., 1996. Transition probability-based indicator geostatistics. Mathematical Geology, **28,** 453-476.

DEUTSCH, C.V. and JOURNEL, A.G., 1998. GSLIB: Geostatistical software Library and User’s Guide. New York, Oxford University Press.

FREEMAN, J.J., 1971. APPROXIMATING DISCRETE PROBABILITY DISTRIBUTIONS. IEEE Transactions on Information Theory, **17,** 491-+.

GOOVAERTS, P., 1994. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. Mathematical Geology, **26,** 389-411.

JOURNEL, A., 1983. Nonparametric estimation of spatial distributions. Mathematical Geology, **15,** 445-468.

KU, H. and KULLBACK, S., 1969. Approximating discrete probability distributions. IEEE Transactions on Information Theory, **15,** 444-447.

LI, W., 2007. Transiograms for characterizing spatial variability of soil classes. Soil Science Society of America Journal **71,** 881-893.

LI, W. and ZHANG, C., 2006. A Generalized Markov Chain Approach for Conditional Simulation of Categorical Variables from Grid Samples. **10,** 651-669.

SAERENS, M. and FOUSS, F., 2004. Yet Another Method for Combining Classifiers Outputs: A Maximum Entropy Approach. In: Multiple Classifier Systems. 82-91.

Table 1: The one dimensional indices of three variables and three facies

k_1	k_2	k_3	$1+\sum_{\ell=1}^{n}(k_\ell-1)\times K^{\ell-1}$	$f(k_\ell,\ell=1,...,n)$
1	1	1	$1+(1-1)*3^{(1-1)}+(1-1)*3^{(2-1)}+(1-1)*3^{(3-1)}$	1
2	1	1	$1+(2-1)*3^{(1-1)}+(1-1)*3^{(2-1)}+(1-1)*3^{(3-1)}$	2
3	1	1	⋮	3
1	2	1	⋮	4
2	2	1	⋮	5
3	2	1	⋮	6
⋮	⋮	⋮	⋮	⋮
2	3	3	$1+(2-1)*3^{(1-1)}+(3-1)*3^{(2-1)}+(3-1)*3^{(3-1)}$	26
3	3	3	$1+(3-1)*3^{(1-1)}+(3-1)*3^{(2-1)}+(3-1)*3^{(3-1)}$	27

Table 2 one numerical example of multivariate probability estimation from the DMPE algorithm

MV index	TRUE MV probability	Bivariate Marginal	estimated MV probability
1	0.0700	0.1390	0.0661
2	0.0350	0.0760	0.0360
3	0.0300	0.0710	0.0329
4	0.0180	0.0950	0.0232
5	0.0410	0.1790	0.0413
6	0.0320	0.0860	0.0265
7	0.0170	0.0750	0.0158
8	0.0150	0.0970	0.0137
9	0.0350	0.1820	0.0376
10	0.0320	0.1050	0.0358
11	0.0340	0.0870	0.0348
12	0.0200	0.0940	0.0154
13	0.0390	0.0910	0.0326
14	0.0890	0.1590	0.0922
15	0.0320	0.1100	0.0351
16	0.0160	0.0970	0.0186
17	0.0360	0.0920	0.0320
18	0.0400	0.1650	0.0415
19	0.0370	0.1350	0.0371
20	0.0260	0.0860	0.0242
21	0.0250	0.0880	0.0267
22	0.0190	0.0910	0.0202
23	0.0490	0.1600	0.0454
24	0.0330	0.1010	0.0354
25	0.0380	0.0670	0.0367
26	0.0350	0.0920	0.0404
27	0.1070	0.1800	0.1029

Table 3 the multivariate probability estimation results from three different approaches

True MV probability	Ku-and-Kullback estimation	DMPE estimation
0.10	0.0998	0.0973
0.10	0.1000	0.1023
0.05	0.0496	0.0495
0.05	0.0493	0.0445
0.00	0.0005	0.0049
0.00	0.0003	0.0027
0.10	0.1001	0.0977
0.05	0.0503	0.0510
0.05	0.0503	0.0504
0.10	0.1000	0.0955
0.00	0.0004	0.0039
0.00	0.0076	0.0066
0.05	0.0495	0.0464
0.05	0.0498	0.0504
0.15	0.1499	0.1499
0.15	0.1496	0.1470

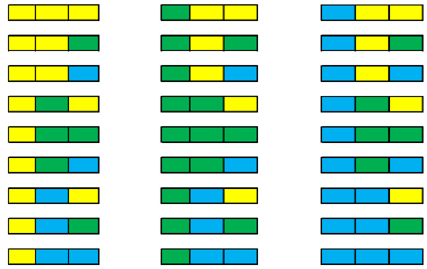

Figure 1 all the states outcomes for 3 categories in 3 grids

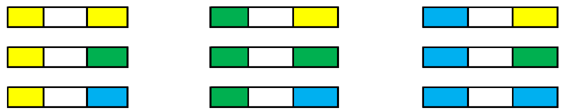

Figure 2 all the states outcomes for one combination of two grids out of three

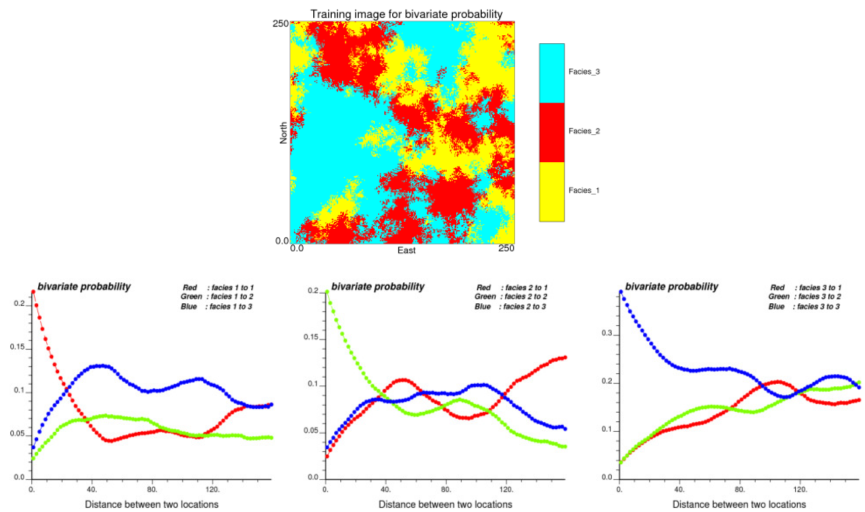

Figure 3 the training image and the inferred bivariate probability diagram

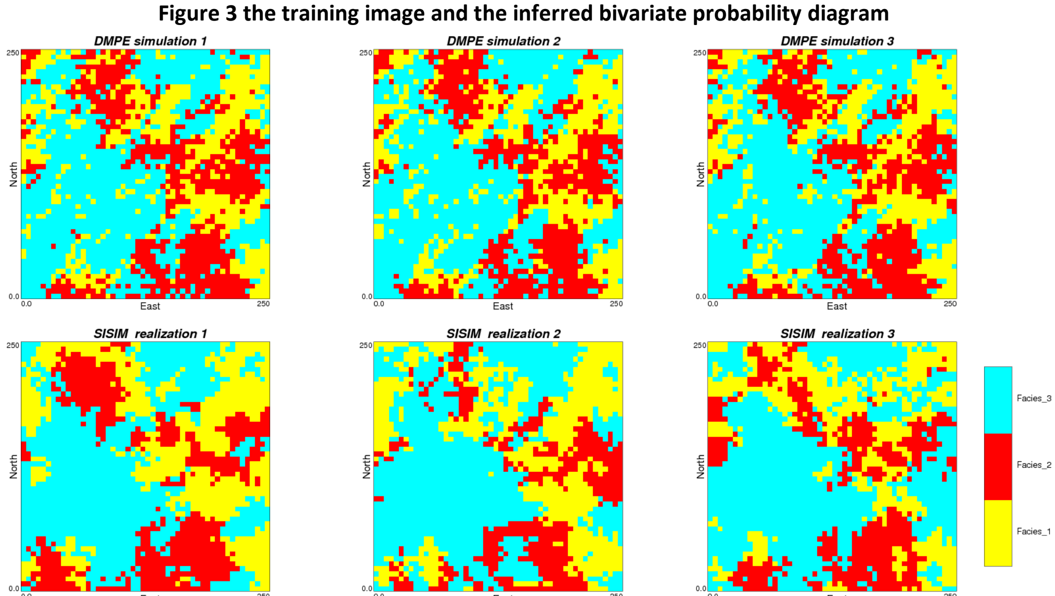

Figure 4 the DMPE simulation results and traditional SISIM simulation results comparison