Application of Multivariate Density Estimation (MDE) to Facies Simulation with Transition Probability Matrices (TPM)

Yupeng Li, Sahyun Hong and Clayton V. Deutsch

In this study, the process of constructing a multivariate probability distribution from specified marginal distribution is addressed. The univariate marginal and bivariate marginal distribution can be obtained from transition probability matrices built from vertical profiles of well data. The 3D conditioning data configurations are transformed to vertical transition probability data configuration during the multivariate probability distribution estimation, which makes it possible to use the vertical transition probability to the 3-D space.

Introduction

In geostatistics, it is very common to build the models of a categorical variable that represents facies or rock types. Although sometimes, deterministical models based on professionals' experiences are used in practice, the geostatistical simulation realizations are being used increasingly for uncertainty quantification. Stochastic modeling algorithms such as sequential indicator simulation (SIS) are widely used to construct these multiple realizations.

Based on the sequential simulation algorithm, SIS is commonly used for categorical variables. The classical SIS algorithm is fast and straightforward because the modeling of the conditional probability distribution at each unsampled location requires the solution of only a single (co)kriging system for each category. Due to the complex features of geological variables, some non-linear geostatistics algorithms have been developed including multiple point geostatistics (Strebelle, 2002). Both the traditional geostatistics (SIS) and multiple geostatistics are based on Bayes Law. A multivariate joint probability is obtained (directly or indirectly) and the conditional probability for the unsampled location given the available data can be obtained using Bayes law. Sequential simulation decomposes the multivariate joint probability by recursive application of Bayes law, while the multiple point statistics set up the multivariate joint probability by scanning a training image (Deutsch, 2002, Caers and Zhang, 2004).

In this paper, a new multivariate probability distribution estimation scheme based on the facies transition probability matrix (TPM) is proposed. The TPM provides the bivariate and univariate marginal distribution for any combination of two categorical variables between any two locations at any lag distance. From this information, specific multivariate probability distributions that satisfy the bivariate and univariate constraints built from the n-conditioning data can be inferred. The estimation process is a kind of iteration based on the transition probability calculated directly from the conditioning data's vertical profile and transformed to 3-D data configuration.

The first section of this paper is an introduction on transition probability and the definition of TPM. Instead of using indicator variograms, transition probabilities are used as the tools to characterize the spatial relationships. In the second section, the construction of a transition probability matrix is introduced. In the third section, the related mathematic relationships between the multivariate probabilities distribution and bivariate marginal distribution are illustrated. Based on that, the multivariate probability distribution estimation (MDE) process is presented. The MDE process presented in this section is a non-linear approach compared with kriging. The fourth section explains the vertical-horizontal transform approach which facilitates 3-D estimation and simulation by this new approach. The final section will present some preliminary results of MDE approach and future work on this new method.

TPM definition

In stochastic theory, the Markov chain is a sequence of random variables (X_1, X_2, ...) with the Markov property, that is, given the present state, the future and past states are independent, which can be written as:

$$P(X_{n+1} = k_{n+1} \mid X_n = k_n, \ldots, X_1 = k_1) = P(X_{n+1} = k_{n+1} \mid X_n = k_n).$$

The possible values of X_t form a countable set $S\ (X_t \in k_i; k_i = 1, \ldots, K, i \leq m)$ called the state space of the chain, where k_i denote mutually exclusive, exhaustively defined states of a stochastic process. If a sequence of states has the Markov property, then every future state is conditionally independent of every prior state. The changes of states are called transitions. The conditioning probability of a state given the previous state $P(X_{n+1} = k_{n+1} \mid X_n = k_n)$ is called the transition probability. The wholly description of the transition probability of a finite state space of S ($k_i, i \leq m$) going from state k_i to state k_j in n steps will form a $m * m$ matrix $T_{i,j}^{(n)}(x, h = n)$ with the ith row and jth column element of $t_{i,j}^{h=n} = P(X_n = k_j \mid X_0 = k_i)$. This matrix is called transition probability matrix (TPM).

Assuming the lithological types at location X_t in the vertical profile of a well will only depend upon the lithological type at the preceding location X_{t-1}, it is recognized as a Markov chain process. Many efforts have been done on the use of Markov chain transition matrix in geology and geostatistics. Most often, the Markov chain transition matrix is used in the vertical profile explanation, sedimentary evolution analysis and stratigraphic sequences simulation. Krumbein has tried molded a transgressive-regressive strand-line deposits using a time-discrete transition matrix to control the lateral shifting(Krumbein, 1968). It is also have been used in soil science to describe the spatial order of different soil cases and vertical spatial change of textural((Li, 1997). Carle and Fogg tried to integrate transition probabilities into the frame of indicator geostatistics for litho-facies simulation (Carle and Fogg, 1996; Carle and Fogg, 1997; Weissmann and Fogg, 1999;Carle, 2000). Elfeki used a kind of coupled markov chain to character the heterogeneity as a non-Gaussian field by multi-dimensional transition probabilities (Elfeki and Dekking, 2001; Elfeki, 2006).

TPM construction

Mainly, the vertical profile is structured as discrete-state Markov chains in two ways. In one way, observations are spaced equally along a vertical profile to yield transition probability matrices. The transitions of the equally spaced rock types at discrete points are counted. Because the same rock type may be observed at successive points, the transition matrix that gives the probability of going from one rock type to another generally has nonzero elements on the main diagonal. The second approach considers only the succession of certain rock types, and because each transition is to a different rock type within the system, the diagonal elements are all zero. In this approach the whole successful thickness of a same rock type, which is one state of Markov chain may recognized from log curve or form outcrop. It is also called an embedded Markov chain.

In this study, the first approach is used to build the TPM. Suppose the whole vertical profile is ***H*** which divided into ***n*** equal segments using an equal segment. The state space in this Markov chain is the facies category set $k_i\,(k_i = 1, 2, .., K)$. Then, in each segment will define a state of a Markov chain and the transition probability of the whole profile will form a TPM.

For example, the total observed number of state k_i followed the state k_j giving the observation interval $\boldsymbol{h}$=**n** is denoted as $\mathbf{n_{i,j}}$ while $\mathbf{n_i}$ is the counted total number of K_i. When the interval is 1, the transition probability from state K_i to state K_j will be:

$$\mathrm{t}_{\mathrm{i,j}}^{\mathrm{h=1}} = n_{i,j}/n_i$$

The probability of a transition from K_1 to K_1, K_2, K_3,...K is given by $t_{1,j}^{h=1}$ (j=1,2,…,m) in the first row and so on and denoted as:

$$T^{h=1} = \begin{pmatrix} t_{i,i}^{h=1} & t_{i,j}^{h=1} & \cdots & t_{i,m}^{h=1} \\ t_{2,1}^{h=1} & \cdots & & t_{2,m}^{h=1} \\ \cdots & & \cdots & \\ t_{m,1}^{h=1} & & \cdots & t_{m,m}^{h=1} \end{pmatrix}$$

Generally, from the account approach, the elements of the TPM ($\forall\ h$) will be:

$$t_{i,j}(x,h) = \left(\frac{n_{i,j}}{n_i}\right) = \left(\frac{n_{i,j} \big/ \sum_{i,j} n_{i,j}}{n_i \big/ \sum_{i,j} n_{i,j}}\right) = \frac{p_{i,j}(x,h)}{p_i(x,h)}$$

Where: $t_{i,j}(x,h)$ is the transition probability of state K_i to state K_j

$n_{i,j}$ is the number of state K_i followed by K_j after **h** steps;

n_i is the row sum of the $n_{i,j}$, $n_i = \sum_j n_{i,j}$;

$\sum_{i,j} n_{i,j}$ is the whole sum of the tally matrix entries;

$p_{i,j}(x,h)$ is the joint probability of two state K_i *and* K_j;

$p_i(x,h)$ is the univariate marginal probability of state K_i

The TPM has to fulfill specific properties:

(1) Its elements are non-negative, $0 \le t_{i,j} \le 1$;

(2) The elements of each row sum up to one, $\sum_{j=1}^{m} t_{i,j}(x,h) = 1$;

Generally,

$$t_{i,j}(x,h) = p\{K_j \text{exit at (x+h)} \,|\, K_i \text{ exit at x}\} = \frac{p\{K_j \text{ exit at (x+h) AND } \mathrm{K}_i \text{ exit at x}\}}{p(K_i \text{ exit at x})}$$

When the process is stationary or homogenous, the transition probability is independent of position x, the transition probability $t_{i,j}(h)$ and the bivariate joint probability $p(\mathbf{h}; k_i, k_j)$ will depend only on the intervals vectors. It shows that the bivariate joint probability:

$$p(\mathbf{h}; k_i, k_j),\ \forall \mathbf{h}; k_i, k_j; i, j = 1, \ldots, K$$

can be calculated from the transition probability $t_{i,j}(x,h)$ as:

$$p(h,k_i,k_j) = p\{k_i \text{exit at (x+h) AND } k_j \text{exit at x}\} = t_{i,j}(h) * p_i(h)$$

The bivariate probability matrix for a particular lag **h** can be also fully defined by its relatively transition probabilities matrix. The sum of all K^2bivariate joint probabilities should be 1. A strong assumption of symmetry would entail that $p(h,k_i,k_j) = p(h,k_j,k_i)$.

TPM curves

For a stationary process, the transition probability will depend on the lag between different observation positions. The transition probability $t_{i,j}(h_i)$ will form a diagram as the h_i increasing from zero to a further distance. For $t_{i,j}(h_i)$ (i=j), that means the states changed to themselves and we can call it direct-transition; if it is $t_{i,j}(h_i)(i \neq j)$, it is called cross-transition which reflects the cross-correlations or inter-states relationship between different states. For example there are 3 rock types in a research well profile. The transition probability matrices curves of rock type 1 changed to rock type 1, 2, and 3 is shown in Figure 1.

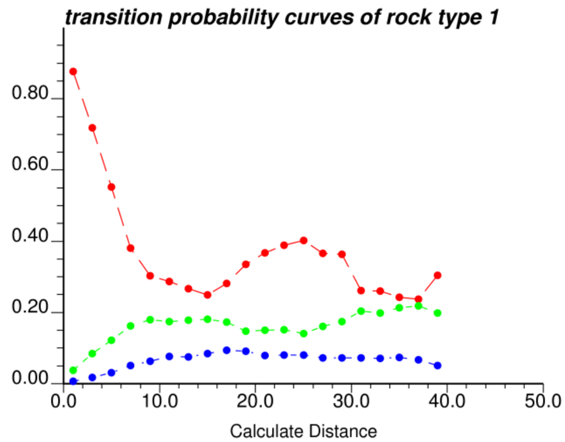

Figure 1 transition probability matrices curves. Red: direct-transition of rock type 1 to 1; Green: cross-transition of rock type 1 to 2; Blue: cross-transition of rock type 1 to 3

As the calculation distance increases, the transition probability of rock type 1 changing to itself is decreasing, with that of changing to rock type 2 and 3 increasing. The same curve can be calculated for the others which compose the whole transition probability matrix curves as shown in figure 2.

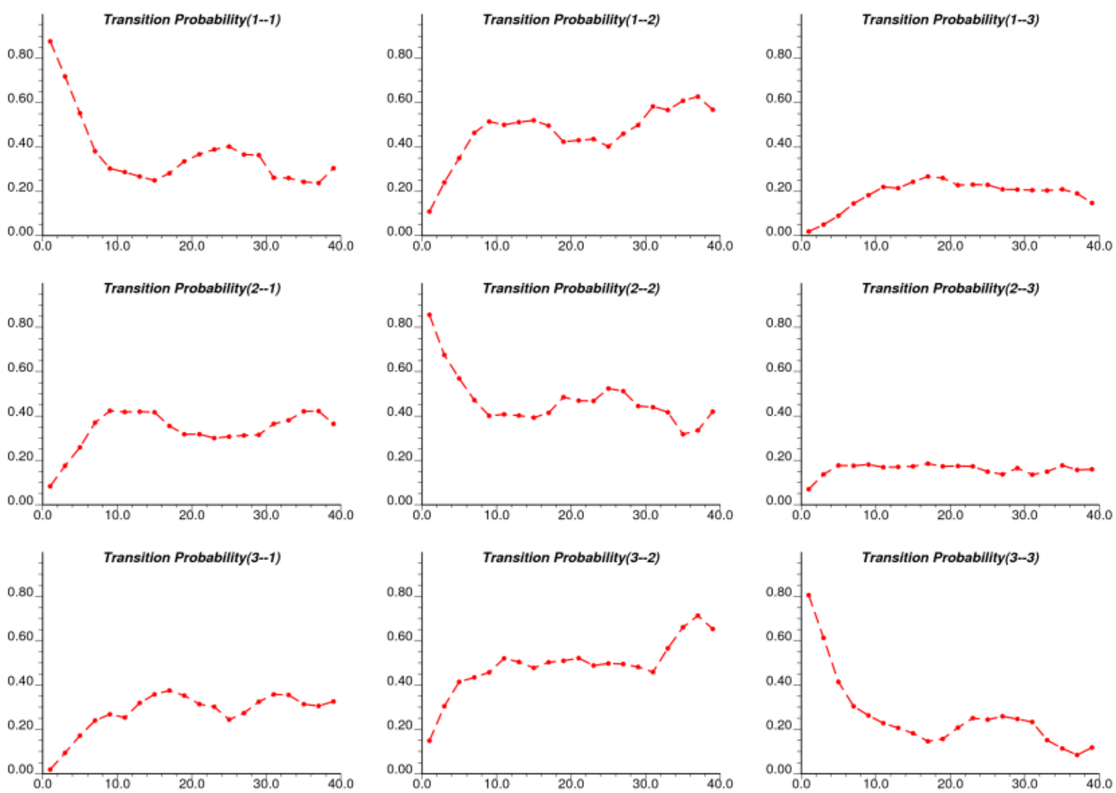

Figure 2 transition probability matrices curves

From the plots in Figure 2, given any known rock type and distance interval for two locations, the transition probability can be calculated. These curves reveal some geological and geostatistical information : (1) as the distance increase, the transition curve will become flat as reach their sill, The TPM will reflect the global univariate probability which can be identified as the percentage of this rock type within the whole section. (2) The TPM will also reflect the relatively bivariate distribution for a particular lag ***h***, in this bivariate distribution the univariate probability at this particular lag ***h*** are also imbedded. (3) The transition probability matrices curves also reflect the spatial juxtaposition information. In Figure 1, during the procedure of rock type 1 decreasing and reached its sill, the other 2 rock types begin to increase and reached their sill with variable probabilities. The rock type 2 has a higher probability to exist than rock type 3 as facies 1 decrease.

Multivariate probability distribution Estimation (MDE) Based on TPM

One of the important problems in geostatistics is finding a proper means to describe spatial dependence. Bivariate probability matrices inferred from TPM can be used as an alternative of variogram to characterize the spatial dependence. Using this full matrix of K^2 bivariate joint probability in multivariate probability distribution estimation is the purpose of this research. Inference from well data leads to a complete specification of the bivariate joint probability matrices for all distance and direction vectors in the vertical profile:

$$p(\mathbf{h}; k_i, k_j), \ \forall \mathbf{h}; i, \ j = 1, \ldots, K$$

Where: $p(\mathbf{h}; k_i, k_j)$ is the bivariate joint probability of rock type k_i and k_j;

k_i and k_j are two rock type in two different location;

h is the observed interval between two different locations.

After the bivariate joint probabilities are inferred, the main concern is the construction of a multivariate joint distribution given these sets of bivariate joint probabilities. Consider the following schematic situation, where the data-data vectors and data-unsampled vectors are all expressed in bivariate joint probability. Of course, the data could be distributed in 3-D space.

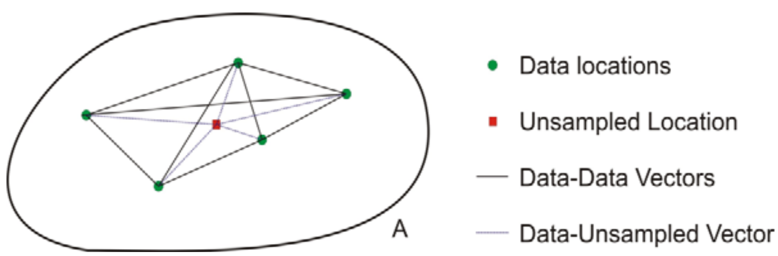

Figure 3 schematic horizontal transition probability data configuration

Assume there are n locations ($u_1 \ldots u_n$), each of the n location has K categories. The n random variables will form a multivariate probability distribution function (pdf) $P_{MV}(u_1, u_2, ..., u_n)$ which is defined as below and represents the probability of a specific configuration of categories $k_i (i = 1, ..., K)$ existing at locations $u_1, u_2, ..., u_n$.

$$P_{MV}(u_1, u_2, ..., u_n) = prob(u_1 \in k_1, u_2 \in k_2, ..., u_n \in k_n); k_i = 1, 2, .., K$$

In this distribution, there are totally K^n possible values. Each of these values occurs with a given frequency which is identified with an index I_{mv} that is:

$$I_{mv} = 1 + \sum_{n=1}^{N} (u_n - 1) * K^{n-1}$$

Where: $I_{mv} = 1, 2, ..., K^n$; u_n is the code of the $\mathbf{n^{th}}$ location of the data configuration that identifies its categories; K is the total number of the categories.

After obtaining this multivariate probability distribution, it is straightforward to calculate the conditioning probability using the Bayes law as below:

$$p(u_0 | u_1, u_2, ..., u_n) = \frac{p_{MV}(u_0, u_1, .., u_n)}{p_{MV}(u_1, .., u_n)} = \frac{P_{MV}(u_0, u_j = k_j)}{\sum_{k_i, k_j = 1}^{K} P_{MV}(u_0 = k_i, u_j = k_j)}$$

$$\text{where: } \mathrm{u_j} = u_0, u_1, ..., u_n;$$
$$k_i, k_j = 1, ..., K; i, j = 1, ..., n$$

The denominator will be the sum of several indices that specified by the conditioning data category values; the nominator is the specific category's multivariate probability. Our challenge is to estimate this multivariate distribution $p_{MV}(u_1, ..., u_n)$ based on the full set of bivariate joint distribution.

Constraints of the multivariate distribution

As stated previously, from the transition probability matrices, we can get the bivariate joint probability $p_{aim}(\mathbf{h}; k', k'')$ of any two facies k' and k''. While if the multivariate probability distribution is known, the bivariate joint probability can be calculated as:

$$p_{cal}(u_{j_1}, u_{j_2}, k', k'') = \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_{j_1}=k'} \cdots \sum_{k_{j_2}=k''} \cdots \sum_{k_n=1}^{K} P_{MV}(u_1, ..., u_n) = P_{k', k''}$$

$$\text{Where: } k_i, k', k'' = 1, 2, ...K; \; i = 1, ...n$$
$$j_1, j_2 = 1, ...n; \; j_1 \neq j_2;$$

The bivariate joint probability $p_{cal}(u_{j1}, u_{j2}, k', k'')$ calculated from the multivariate probability distribution should be equal those obtained from the transition probability matrix $p_{aim}(\mathbf{h}; k', k'')$. This will impose $(\mathbf{n} * (\mathbf{n} - \mathbf{1}) * \mathbf{k^2})$ constraints on the multivariate probability distribution.

The order relationship of the multivariate distribution will compose another constraint to this multivariate probability distribution.

$$\sum_{i=1}^{K^n} p_{MV}(u_1, ..., u_n) = 1 \quad u_1, ..., u_n = 1, ..., K$$

Based on those two constraints, an iteration approach is adopted to modify an initial multivariate probability distribution to satisfy those two constraints. The steps are:

Step 1: The initial value of the multivariate probability values come from assumption that the facies at each locations are independent. Under the independent assuming, the multivariate distribution constrained only to the univariate probabilities could be written as follows:

$$p_{MV}(\mathbf{u}_1 = k_1, ..., \mathbf{u}_n = k_n) = \prod_{j=1}^{n} p_{k_j} \qquad k_1, ..., k_n = 1, ..., K$$

Step 2: The initial distributions are modified by the constrained of transition probability which expressed as a bivariate joint probability between every two locations. After modified by bivariate joint probability, a new multivariate probability distribution is:

$$p^*_{MV}(\mathbf{u}_0 = k_0, \mathbf{u}_1 = k_1, \ldots, \mathbf{u}_n = k_n)$$

Step 3: Go to the bivariate joint probability constraints using p^*_{MV} as the new initial value for the distribution until there are no notable changes in the conditioning probability calculated from the multivariate probability distribution.

The iteration modifying process is shown in figure 4.

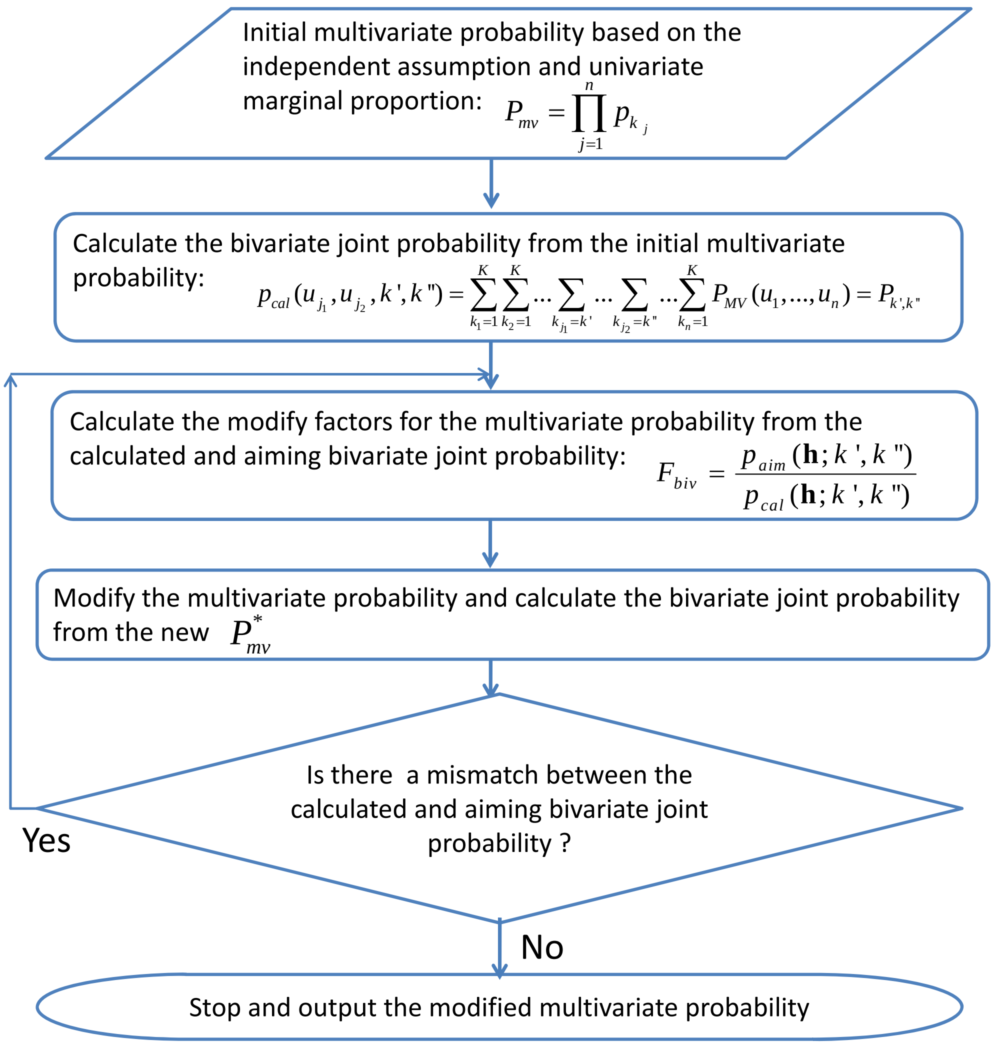

Figure 4 the modifying process of MDE

Transformation of TPM to any Spatial Direction

Usually, in the vertical profile of the well data or outcrop, the data have a higher density which can build the transition probability matrices more easily. While for estimation and simulation, the grid needs to be simulated in a 3D space.

Geological research have clear seen that sedimentary facies show vertical sequence superposition and the vertical progression of facies will reflect lateral facies changes. Sedimentary environments that started out side-by-side will end up overlapping one another over time due to transgressions and regressions. The result is a vertical sequence of facies mirrors the original lateral distribution of sedimentary environments (http://www.wvup.edu/ecrisp/g103lecstratpinciples.html). It is called Walther's law named after a German geologist. Later, more advanced theories such as sequence stratigraphy theories are used to correlate the vertical profile with the lateral shifts. They all provide clues to explain how the facies and surrounding deposits change and shift laterally from the vertical facies deposition stack pattern as earth's surface undergoes changes. As an example, in a single transgression cycle (a relative rise in sea level resulting deposition of marine strata over terrestrial strata), the vertical facies stacking pattern will related to horizontal pattern as shown in Figure 5.

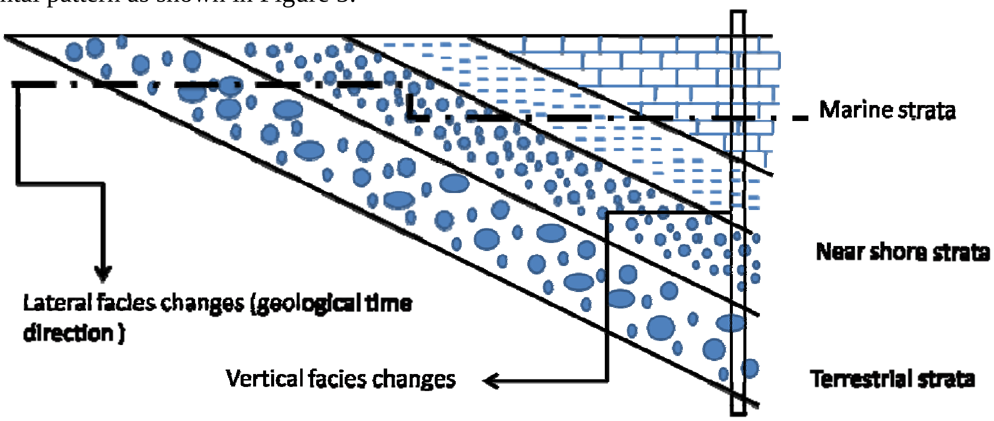

Figure 5 Walther's Law in a single transgression cycle, which shows that facies vary in an analogous manner both horizontally and vertically

After the transition probability are calculated in vertical profile, based on Walther's Law, a kind of horizontal to vertical transformation ratio can be used to provide the transition probability information for any lateral horizontal direction. Now assuming in a basin marginal as shown in Figure 6, direction along location ***a*** to location ***b*** is the direction toward shoreface, where the juxtaposition will become more sandy. From location ***a*** to location ***f*** is the direction leaving the shoreface and toward the offshore. The juxtaposition in this direction will become more mudy. The shore line is nearly parallel to the direction of ***a*** to ***d***, and we believe those locations are all in a single sedimentary sequence cycle.

Distance vector ***a-c*** is any distance vector which can decompose into three vectors: along the direction of ***a-b,*** direction of ***a-d*** and the vertical direction. Given the anisotropy ratio along those three directions as: , the effectively distance in vertical transition probability matrix of ***a-c*** will be:

$$h_{effect}(c) = \sqrt{(\frac{h_{main}}{a_{main}})^2 + (\frac{h_{\min}}{a_{\min}})^2 + (\frac{h_{vert}}{a_{vert}})^2}$$

Where:

Anisotropy ratio along the main difference direction;

Distance of unsampled location departs from the sampled location along the main difference direction;

Anisotropy ratio along the minor difference direction;

Distance of unsampled location departs from the sampled location along the main difference direction;

Anisotropy ratio along the vertical direction;

Distance of unsampled location departs from the sampled location along the vertical direction;

Effective distance along the sampled location's vertical direction;

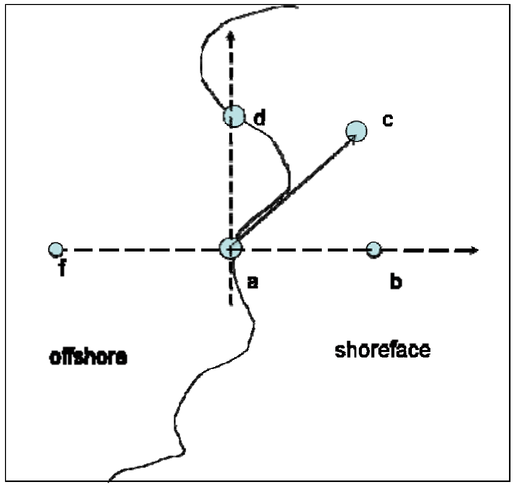

Figure 6 Estimation well point distribution pattern

Where: the direction and the value of $a_{main,}a_{\min,}$ and a_{vert} will depend on the geological understanding. After the effectively distance vector $h_{effect}(c)$ calculation, the TPM of location ***c*** can be read from the vertical profile location ***a***. By this approach, any horizontal data configuration can be transformed into vertical data configuration which the bivariate model has been built from the vertical profile data as shown in Figure 7.

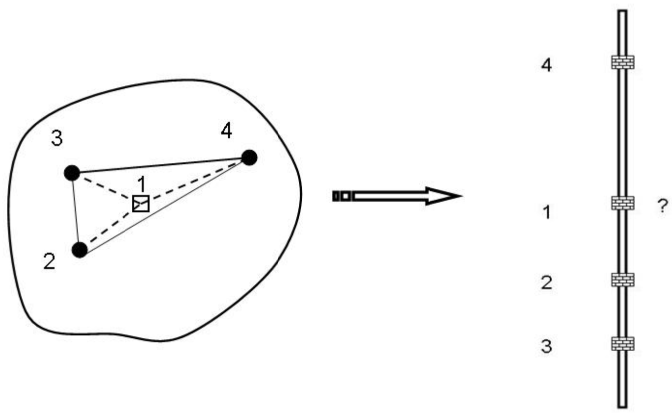

Figure 7 Horizontal data configuration to vertical data configuration transform

In the vertical direction, we have finer data information; the vertical transition probability will provide any distance transition probability for the horizontal direction.

Some test case and estimation/simulation example

Based on the previous multivariate probability distribution estimation algorithm (the program details are explained later in this report), one small 3-category example is used for results illustration. In most cases, the categories in reservoir can separated into 3 categories according their qualities: non-net, media and net sand. The data set location map is shown in upper left of Figure 8. The variogram modes for those three categories are also shown in upper right of Figure 8. The transition probability matrix for those three categories comes from a true well data. The totally lag count number is 40, and then the transition probability matrices is shown in bottom of Figure 8.

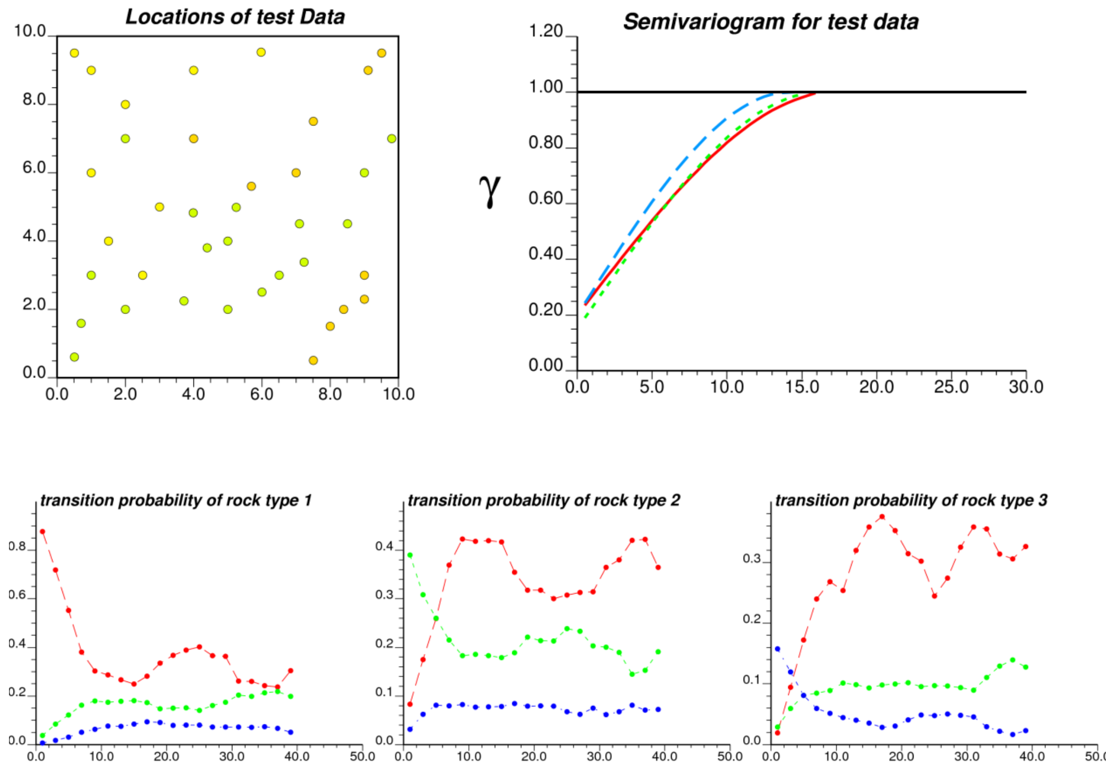

Figure 8 the test data location map, variogram models and transition probability matrices for the test case

The research grid is 10 by 10 grid nodes (100 totals). Based on that information, the estimation and simulation are done. The estimation results including IK and MDE results are shown in Figure 9.

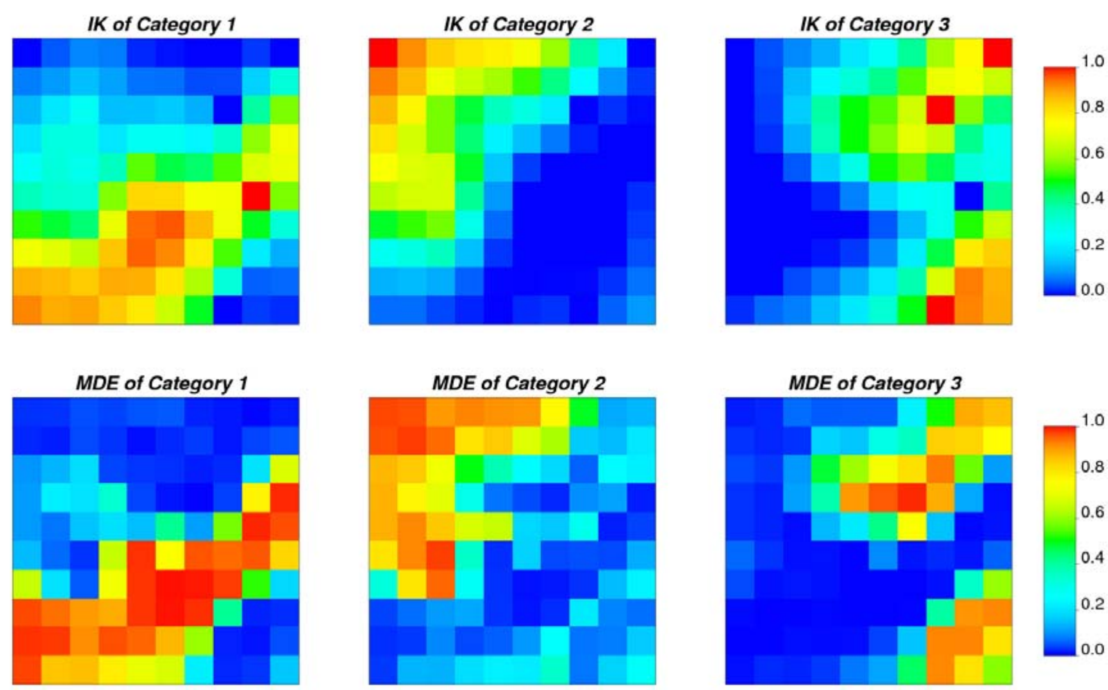

Figure 9 IK and MDE estimation results

For simulation, three kinds of simulation algorithm are used. One is the traditional *sisim*. While for the *blocksis*, the results are cleaned with a light clean option. The last one is simulation with MDE. The results are shown in Figure 10.

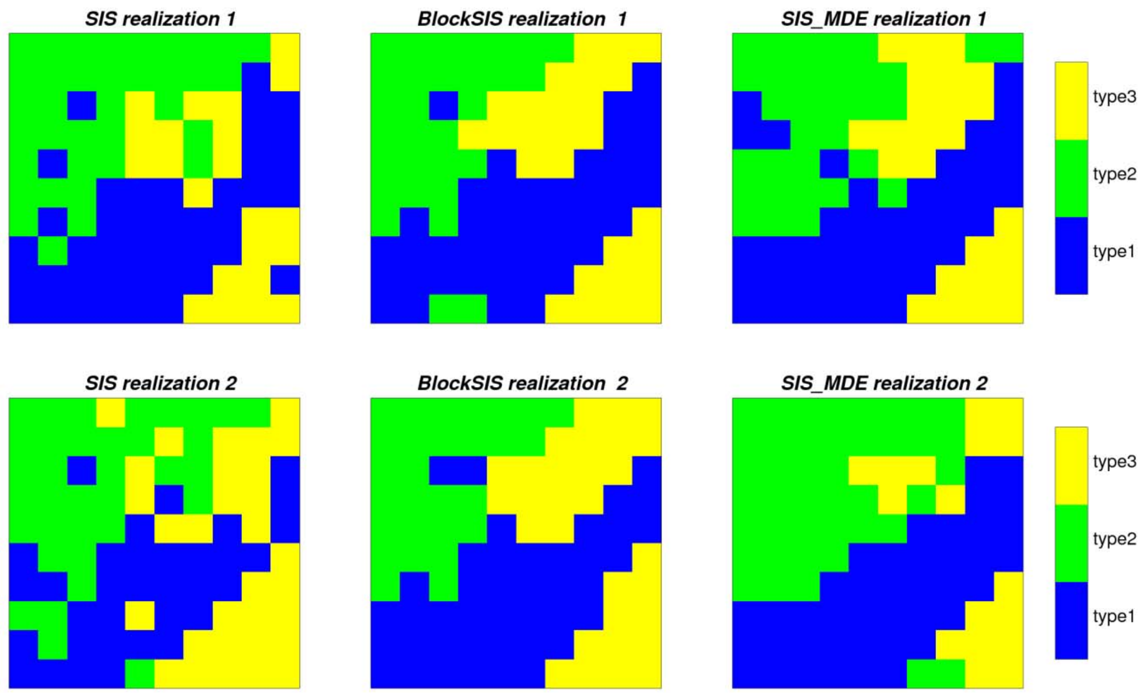

Figure 10 SIS realizations with different algorithm

Conclusions and Future Work

The MDE based on TPM and the horizontal-vertical transition probabilities transformation approach will provide a new path for categorical variable estimation and simulation. No kriging is involved and yet a completely consistent multivariate distribution is predicted for each unsampled location constrained to all available information. In this approach, the bivariate joint probabilities are used directly; there are no such order relation deviation problems as occurs in traditional indicator kriging. The bivariate joint probability is easy to integrate with univariate probability information, which will provide a platform to account for secondary data inferred from the available site-specific observation. This could also be generalized in future research. Because the multivariate probability has a large dimension, and only bivariate joint probabilities constraints are used in each iteration, the iteration convergence speed may be slow and CPU intensive. Future research will address the speed of the algorithm.

Acknowledgements

The authors would like to acknowledge the support of the sponsor companies of the Centre for Computational Geostatistics.

References

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

—, 1997. Modeling Spatial Variability with One and Multidimensional Continuous-Lag Markov Chains. *Mathematical Geology*, **29,** 891-918.

Carr, T.R., 1982. Log-linear models, Markov chains and cyclic sedimentation. *Journal of Sedimentary Research*, **52,** 905-912.

Caers, J., Direct sequential indicator simulation. In: Kleingeld, W.J., Krige, D.G. (Eds.), Proceedings of the Geostatistics 2000. Cape Town, South Africa, pp. 39–48.

Caers, J., Zhang, T.,2004, Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models. In Grammer, G.M., et al. (Eds), integration of Outcrop and Modern Analogs in Reservoir Modeling. Tulsa, OK, AAPG Memoirs, pp. 383–394.

Dacey, M. and Krumbein, W., 1970. Markovian models in stratigraphic analysis. *Mathematical Geology*, **2,** 175-191.

Deutsch, C.V., Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data, Ph.D. Thesis, Stanford University, 1992

Deutsch, C. V., and Journel, A.G., 1998, GSLIB: Geostatistical Software Library, Oxford University Press, New York.

Deutsch, C. V., 2002, Geostatistical Reservoir Modeling, Oxford University Press, New York.

Deutsch C. V., Julian M. Ortiz, and Chad Neufeld; New software for fitting indicator covariance and for indicator Kriging/simulation; Centre for Computational Geostatistics Annual Report 7 (paper-401),2005

Elfeki, A. and Dekking, M., 2001. A Markov Chain Model for Subsurface Characterization: Theory and Applications. Mathematical Geology, 33, 569-589.

Elfeki, A.M.M., 2006. Reducing concentration uncertainty using the coupled Markov chain approach. *Journal of Hydrology*, **317,** 1-16.

Fogg, G.E., Noyes, C.D. and Carle, S.F., 1998. Geologically based model of heterogeneous hydraulic conductivity in an alluvial setting. *Hydrogeology Journal*, **6,** 131-143.

Hattori, I., 1973. Mathematical analysis to discriminate two types of sandstone-shale alternations, 331-345.

Krumbein, W.C., 1968. STATISTICAL MODELS IN SEDIMENTOLOGY1. *Sedimentology*, **10,** 7-23.

Krumbein, W.C. and Dacey, M.F., 1969. Markov chains and embedded Markov chains in geology. *Mathematical Geology*, **1,** 79-96.

Kulatilake, P.H.S.W., 1987. Modelling of cyclical stratigraphy using Markov chains. *Geotechnical and Geological Engineering*, **5,** 121-130.

Lee, S.-Y., Carle, S.F. and Fogg, G.E., 2007. Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. *Advances in Water Resources*, **30,** 1914-1932.

Leuangthong, O., and Deutsch, C.V., "Stepwise Conditional Transformation for Simulation of Multiple Variables", *Mathematical Geology*, 35 (2), February 2003, pp. 155-173.

Li, W., 2007. Markov Chain Random Fields for Estimation of Categorical Variables. *Mathematical Geology*, **39,** 321-335.

Maji, R., Sudicky, E.A., Panday, S. and Teutsch, G., 2006. Transition Probability/Markov Chain Analyses of DNAPL Source Zones and Plumes, 853-863.

Miall, A.D., 1973. Markov chain analysis applied to an ancient alluvial plain succession, sedimentology, **20**, 347-364.

S. Strebelle, Conditional simulation of complex geological structures using multiple-point statistics, Mathematical Geology **34** (7) (2002), pp. 1161–1168.

Weissmann, G.S. and Fogg, G.E., 1999. Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *Journal of Hydrology*, **226,** 48-65.