


# Spatial hidden Markov chain models for estimation of petroleum reservoir categorical variables

Xiang Huang<sup>1</sup>  · Jie Li<sup>2</sup> · Yuru Liang<sup>3</sup> · Zhizhong Wang<sup>1</sup> · Jianhua Guo<sup>2</sup> · Peng Jiao<sup>2</sup>

Received: 17 October 2015 / Accepted: 2 May 2016 / Published online: 9 May 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Indicator variograms and transition probabilities are used to measure spatial continuity of petroleum reservoir categorical variables. Variogram-based Kriging variants are symmetric geostatistical methods, which cannot completely capture the complex reservoir spatial heterogeneity structure. The asymmetric spatial Markov chain (SMC) approaches employ transition probabilities to incorporate proportion, length and juxtaposition relation information in subsurface reservoir structures. Secondary data in petroleum geology, however, cannot be reasonably aggregated. We propose a spatial hidden Markov chain (SHMC) model to tackle these issues. This method integrates well data and seismic data by using Viterbi algorithm for reservoir forecasting. The classified sonic impedance is used as auxiliary data, directly in some kind of Bayesian updating process via a hidden Markov model. The SMC embedded in SHMC has been redefined according to first-order neighborhood with different lag in three-dimensional space. Compared with traditional SMC in Markov chain random field theory, the SHMC method performs better in prediction accuracy and reflecting the geological sedimentation process by integrating auxiliary information.

**Keywords** Hidden Markov model · Posterior probability · Reservoir simulation · Sonic impedance · Viterbi algorithm

## Introduction

Markov process was firstly introduced by Russian mathematician, Andrei Andreyevich Markov, in 1907. Markov-based algorithm, like Markov chain Monte Carlo (MCMC) simulation, has recently been used to quantify uncertainty in infill well placement in the field of petroleum exploration and production (Arinkoola et al. 2015). Spatial Markov chain (SMC) models have also been widely adopted in petroleum reservoir to characterize the spatial heterogeneity of categorical variables through the conditional probabilities (transition probabilities) from different directions (Carle and Fogg 1997; Weissmann and Fogg 1999). At present, there are two kinds of different independent assumptions to simplify the conditional probability of SMC models: one is full independence assumption; the other is conditional independence assumption. The full independence assumption is defined by Elfeki and Dekking (2001), and the corresponding conditional probability formulas are proposed by Elfeki and Dekking (2001) and developed by Li et al. (2012). A spatial Markov chain with full independence assumption consists of several one-dimensional Markov chains, which are forced to move to the same location with equal states. The full independent assumption caused the small class underestimation problem. This method is feasible only for enough conditional data. The conditional independence assumption can be found in Pickard random field in cardinal directions (Pickard 1980), and its general definition is suggested by Li (2007), i.e., given a cell, its nearest neighboring states are conditionally independent. The general conditional probability formulas are given by Li (2007), which do not have the small class underestimation problem.

An SMC is actually a dimensionality reduction process where its multi-dimensional conditional probabilities are

✉ Xiang Huang  
huangxiang@csu.edu.cn

<sup>1</sup> Department of Statistics, Central South University, Changsha 410012, Hunan, China

<sup>2</sup> School of Geosciences and Info-Physics, Central South University, Changsha 410083, Hunan, China

<sup>3</sup> Institute for Basic Research, Shandong Yingcai University, Jinan 250104, Shandong, China

expressed as multiple one-dimensional transition probabilities. The transition probabilities of reservoir categorical variables, such as lithofacies, can be estimated from well data. The vertical transition probabilities can be estimated by the vertical transition tallies from well logs. The transition probabilities in other directions can be estimated by the Walther’s law (Li et al. 2012). Most traditional geostatistical models, like Markov chain random field (MCRF), use well data only and make prediction based on SMC, which results in a relatively low prediction accuracy (Huang et al. 2016a). Huang et al. (2016b) introduced a beta-transformed Bayesian updating model to boost the classification accuracy of category random field. Auxiliary information, however, has not been taken into consideration. To make use of secondary data, such as geophysical well logs, Eidsvik et al. (2004) used hidden Markov chains for estimation of geological attributes. The hidden Markov chain uses Dirichlet prior distributions for the Markov transition probabilities between rock types. Li et al. (2010) developed the Markov chain models by integrating multi-scale information, such as logging, core data and seismic data. In the remote-sensing area, Li et al. (2015) introduced a Bayesian MCRF cosimulation method for improving land cover classification accuracy. We propose a single spatial hidden Markov chain (SHMC), which improves the accuracy of reservoir modeling by integrating geological conceptual data with well data.

## Review of Markov models

### Markov mesh model

A petroleum reservoir grid is a finite, regular grid in one to three dimensions, and its gridding cells are indexed by a positive integer  $s$ , where  $s$  takes on values in  $S = \{1, 2, \dots, n\}$ . All cell states  $F = \{F_1, F_2, \dots, F_n\}$  can be regarded as a family of random category variables defined on the set  $S$ ; each random variable  $F_s$  takes a state value  $f_s$  in the state set  $\Omega = \{1, 2, \dots, m\}$ . If all cell states  $F_1, F_2, \dots, F_n$  follow a sequential path, it is defined as a spatial stochastic sequence. A set of reservoir category variables  $F$  can be considered as a Markov random field or a Gibbs random field, and its joint probability (likelihood function) generally takes the following form (Tjelmeland and Besag 1998; Salomão and Remacre 2001)

$$\Pr(f) = \frac{\exp\left\{-\sum_{s=1}^n \sum_{j \in \eta_s} \Psi(f_s, f_j)\right\}}{\sum_f \exp\left\{-\sum_{s=1}^n \sum_{j \in \eta_s} \Psi(f_s, f_j)\right\}} \quad (1)$$

where  $\eta_s$  is a set of cells which is adjacent to  $s$ ;  $\Psi(f_s, f_j)$  denotes the relationship between cell  $s$  and cell  $j$ ;  $f =$

$\{f_1, f_2, \dots, f_n\}$  is a configuration of  $F$ , corresponding to a realization of the field. The use of Eq. (1) for the simulation of reservoir category variables is theoretically feasible, but it is actually limited by the highly time consuming in computation. By the best-known classical approximation, Eq. (1) can be simplified as Blake et al. (2011) suggested

$$\Pr(f) = \prod_{s=1}^n \Pr(f_s | f_{\eta_s}) \quad (2)$$

where  $f_{\eta_s} = \{f_r | r \in \eta_s\}$  stands for the set of state values at the cells neighboring  $s$ .

Markov mesh models (Stien and Kolbjørnsen 2011) are fully specified through the conditional probabilities in RHS of Eq. (2) as

$$\Pr(f_s | f_{\eta_s}) = \Pr(f_s | f_{s-1}, f_{s_1}, f_{s_2}, \dots, f_{s_l}) \quad (3)$$

where  $s_1, s_2, \dots, s_l$  is its nearest known locations of current cell  $s$  in different directions;  $s - 1$  is always the start cell of the Markov chain to the unknown cell  $s$ , which is to be estimated. The probabilities in Eq. (3) are defined through logit link functions in generalized linear models, and Markov mesh can use larger cliques or neighborhood to capture complex interclass relationships. Recently, Stien and Kolbjørnsen (2011) proposed the method of a fast estimation through iterated weighted least squares and fast simulation through a unilateral path. Kolbjørnsen et al. (2014) recommended using multiple grids in Markov mesh facies modeling, which is typically ten times faster than that of creating one SNESIM realization. Although Markov mesh model is widely used in geoscience, the parameter estimation and iteration process are annoying.

### Spatial Markov chain models

Spatial Markov chain models use the full independence assumption and the conditional independence assumption to define the conditional probability for simplifying the complex computation in Eq. (3). It is actually a dimensionality reduction process where its conditional probabilities  $\Pr(f_s | f_{\eta_s})$  are expressed as multiple one-dimensional transition probabilities from different directions. The spatial Markov chain can be constructed by  $l + 1$  one-dimensional Markov chains together, but these one-dimensional chains are forced to move to the same location with equal states under the full independence assumption. Then, the conditional probabilities in Eq. (3) can be expressed as

$$\Pr(f_s | f_{s-1}, f_{s_1}, \dots, f_{s_l}) = \frac{p_{f_{s-1}f_s}^1 p_{f_{s_1}f_s}^1 \dots p_{f_{s_l}f_s}^l}{\sum_{f_s} p_{f_{s-1}f_s}^1 p_{f_{s_1}f_s}^1 \dots p_{f_{s_l}f_s}^l} \quad (4)$$

where  $p_{f_r f_s}^r$  denotes a transition probability in the  $r$ th direction from state  $f_{s_r}$  to  $f_s$  and  $p_{f_{s-1}f_s}$  denotes a transition

probability along moving direction of the spatial Markov chain from state  $f_{s-1}$  to  $f_s$ . We can derive the conditional probabilities of two-dimensional Markov chain model (Elfeki and Dekking 2001) and three-dimensional Markov chain model (Li et al. 2012) from Eq. (4). Using the conditional independence assumption, Li (2007) gives the general expression of the conditional probability formula in Eq. (3) at any location  $s$  as

$$\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l}) = \frac{p_{f_s-1f_s} p_{f_s f_{s_1}}^1 \dots p_{f_s f_{s_l}}^l}{\sum_{f_s} p_{f_s-1f_s} p_{f_s f_{s_1}}^1 \dots p_{f_s f_{s_l}}^l} \quad (5)$$

where  $p_{f_s f_{s_r}}^r$  denotes a transition probability in the  $r$ th direction from state  $f_s$  to  $f_{s_r}$  and  $p_{f_{s-1}f_s}$  denotes a transition probability along moving direction of the spatial Markov chain from state  $f_{s-1}$  to  $f_s$ .

Generally speaking, the difference between spatial Markov chain model and Markov mesh model is that the latter uses directly the local conditional probabilities in Eq. (3) or the joint probability in Eq. (2), and spatial Markov chain models use multiple one-dimensional transition probabilities or simplified formulas of the local conditional probabilities in Eq. (3). A spatial Markov chain model may be viewed as a special case of Markov mesh models, whereas a Markov mesh model is an extension of spatial Markov chain models, called a generalized spatial Markov chain model.

### Spatial hidden Markov chain model

A spatial hidden Markov chain (SHMC), a combination of SMC and hidden Markov model (HMM), is a double random sequence process consisting of a Markov chain and a spatial stochastic sequence. It can make good use of information from well data and auxiliary data. The SHMC is an extension of SMC. It is better able to capture inter-class dependency relationships (neighboring relationships, cross-correlations, directional asymmetries) among hidden variables. A spatial Markov chain  $F = \{F_1, F_2, \dots, F_n\}$  of reservoir categorical variables is characterized by its states and conditional probabilities through Eqs. (4) or (5), and the model is particularly useful as a prior model. The states of the chain except the wells are unobservable, therefore “hidden.” A stochastic sequence  $W = (W_1, W_2, \dots, W_n)$  of reservoir categorical variables is from auxiliary data, and its observed values are denoted by  $w = (w_1, w_2, \dots, w_n)$ .

### Definition based on Bayes theory

A spatial hidden Markov model uses the posterior probability distribution for modeling reservoir categorical variables and the distribution of the possible states  $F$ . Given the

observations  $w$ , the posterior probability is computed by using the following formula

$$\Pr(f|w) = \frac{\Pr(w|f)\Pr(f)}{\sum_f \Pr(w|f)\Pr(f)} \quad (6)$$

Using Eq. (6), the local conditional probabilities are written as

$$\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l}, w_s) = \frac{\Pr(w_s|f_s)\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l})}{\sum_{f_s} \Pr(w_s|f_s)\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l})} \quad (7)$$

where  $\Pr(f)$  is the prior probability, which is estimated from well data;  $\Pr(w|f)$  is conditional probabilities of the observations  $w$  for  $f$  fixed, i.e., a likelihood item;  $\Pr(w) = \sum_f \Pr(w|f)\Pr(f)$  is the probability of  $W$ , which is a normalization constant when  $w$  is given. We call unobservable  $f$  “true states” and  $w$  “observed values.” The right side of Eqs. (6) or (7) has been widely used since Thomas Bayes (1764) and Pierre–Simon Laplace (1774) introduced Bayesian statistics, but it is not found in petroleum reservoir hidden Markov application. To simulate reservoir categorical variables using  $\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l}, w_s)$ , we need to estimate the conditional probability  $\Pr(w_s|f_s)$  and compute the local conditional probability  $\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l})$  in the right of Eq. (7).

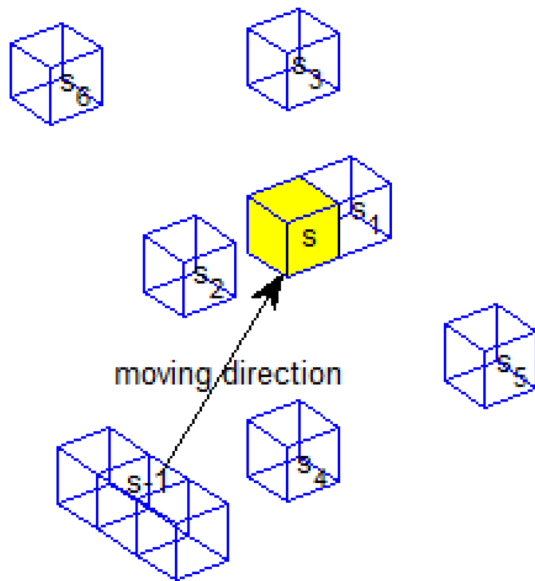
### Specifying the prior conditional probability

We use Eq. (8) to define the prior conditional probabilities; the formula is given as follows

$$\Pr(f_s|f_{s-1}, f_{s_1}, \dots, f_{s_l}) = \frac{p_{f_s-1f_s} p_{f_s f_{s_1}}^{(1)} \dots p_{f_s f_{s_l}}^{(l)}}{\sum_{f_s} p_{f_s-1f_s} p_{f_s f_{s_1}}^{(1)} \dots p_{f_s f_{s_l}}^{(l)}} \quad (8)$$

where  $s_1, s_2, \dots, s_l$  is its nearest known locations of current cell  $s$  in different directions;  $s - 1$  is always the start cell of the Markov chain to the unknown cell  $s$ , which is to be estimated; the superscript (1), (2), ..., (l) indicates the different lag  $h$ .

Thus, we have redefined SMC illustrated in Fig. 1. The prior conditional probabilities can be computed with Eq. (8), where  $p_{f_s f_{s_l}}^{(l)}$  is given in transition probability function. Obtaining the local conditional probabilities requires to calculate  $\Pr(w_s|f_s)$ . The state value  $f_s$  is the “true value” and unobservable except the well data, and the state value  $w$  is regarded as “observation value.” It is noted that  $\Pr(w_s|f_s)$  is essentially the likelihood or the forward model relating facies to sonic impedance; it is not the prior geologic concept, though of course the geology helps to pick the right rock physics and seismic model to relate facies to impedance.



**Fig. 1** SMC defined by first-order 3-D neighborhood with different lag, cell  $s$  is to be studied

## Case study

### Data set

The data we used for our research are gathered from Tahe area of the Tarim Basin in Xinjiang Uygur Autonomous Region, China. Tahe oil field, located in Xayar uplift, north of Tarim basin (Fig. 2), up to now is one of the greatest domestic discoveries in the Paleozoic carbonate rock series. There are two extensive unconformities developed in this area. The Carboniferous clastic rocks directly overlie on the carbonate rocks of Ordovician and underlie the Permian pyroclastic rocks or Triassic formation (Fig. 3). In view of achievements in the carbonate formation of Ordovician, as the seal of it, the Carboniferous (T50–T56), which belongs to the same petroleum system as Ordovician, also shows its exploration potential of lithological reservoir. The purpose layer is located at the depth between 5200 and 5300 m, developed as part of the second formation of the lower Carboniferous (Kalashayi formation). There are three major lithofacies in this work area: mudstone, sandstone and conglomerate. The conglomerate is relatively low in content. We have got four wells' log data with 509 samples in the three-dimensional space, just as shown in Fig. 4. Three wells are located in the corners of this work area; another well is located inside. The distance in east–west direction of the two wells is 900 and 1200 m in south–north direction, the simulated space is split into a  $30 \times 40 \times 50$  grid system, and each cell is a  $30 \text{ m} \times 30 \text{ m} \times 2 \text{ m}$  cuboid.

## Impedance partition

Figure 5 illustrates some basic descriptive statistics of sonic impedance, such as sample size, mean, variance. The impedance will be regarded as observed value  $w$  and will be divided into two classes: strong and weak. By analyzing log and core data, we choose the impedance median 8315.48 as the threshold. The impedance is regarded as strong if it is greater than 8315.48 and weak when it is less than the threshold. By doing so, we derive the emission matrix (Table 1) and the emission probability (Table 2). The initial proportion of each reservoir categorical variable can be computed from Table 1. Mudstone is 62.87 %, sandstone is 31.24 %, and conglomerate is 5.89 %, respectively. The impedance of each class is depicted in Fig. 6. By analyzing Table 2 and Fig. 6, we may find that the impedance of mudstone tends to be stronger than sandstone and conglomerate.

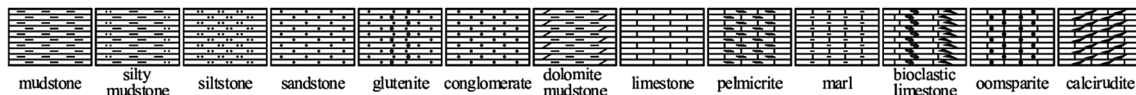
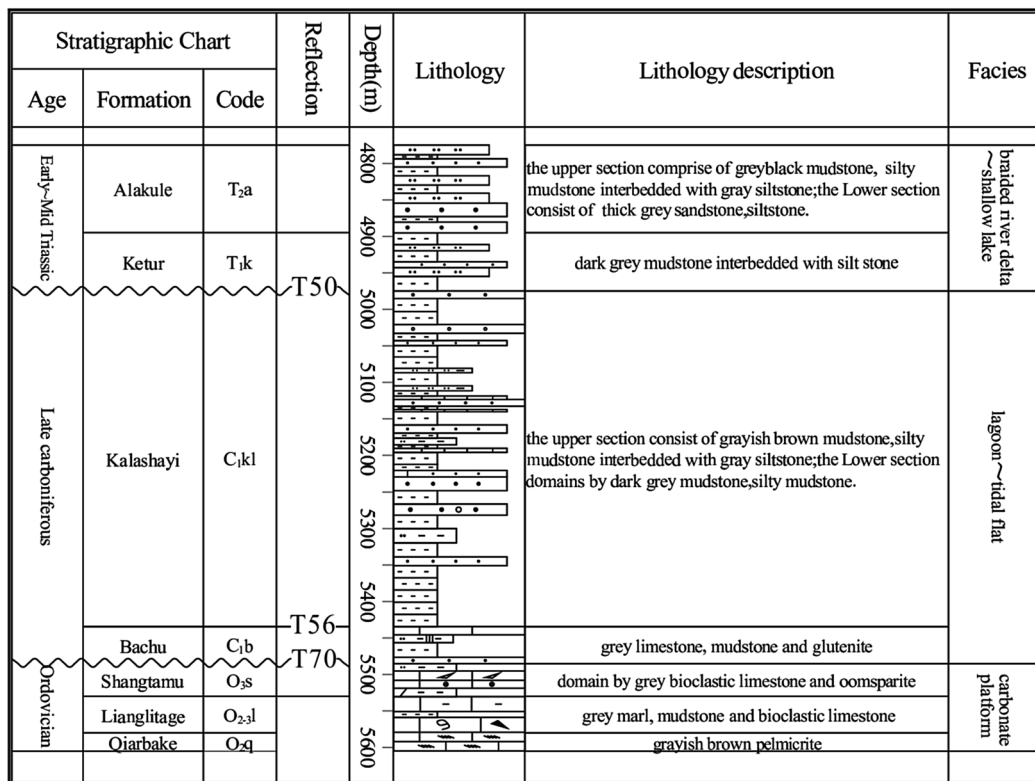
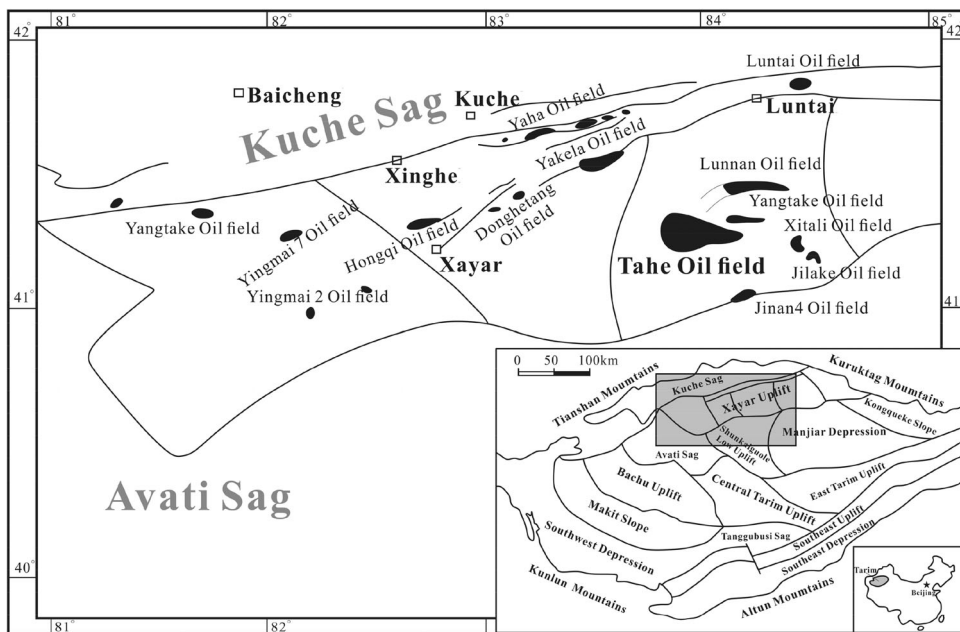
## Transiogram models

The magnitude of the transition probability depends on a sampling interval, i.e., the transition probability is a non-linear function of the sampling intervals (Carle and Fogg 1997). By increasing sampling interval, the transition probability forms a transition probability function (also called “transiogram”), which is regarded as a measure of spatial continuity (Li 2006). Experimental transiograms are estimated from the 509 points and fitted by exponential models. The fitted transiogram models are used for simulations. Because raster data are used, the lag  $h$  represents as the number of pixels (i.e., grid units), not the exact distance. Figure 7 illustrates the experimental auto-/cross-transiograms and their fitting models. It can be seen that most of these experimental transiograms can be approximately fitted by an exponential model. We also find that some experimental transiograms have apparent fluctuations that are difficult to fit using the basic model, such as  $p_{22}$  and  $p_{23}$ . This may be caused by the insufficiency of observed data and the non-Markovian effect of the real data. Fitted transiogram models capture only part of the features of experimental transiograms, depending on the complexity of the mathematical models used (Li 2007). Using composite hole-effect models (Ma and Jones 2001) may capture more details, such as periodicities, of experimental transiograms.

## Simulation results

The SHMC can be determined by initial probabilities (prior probabilities)  $C$ , transition matrix  $A$  and emission matrix  $B$ . The states sequence depends on  $C$  and  $A$ , while the

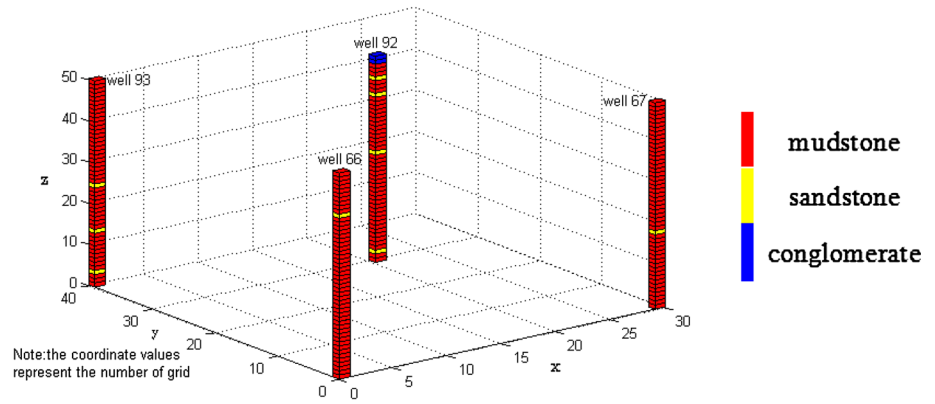
**Fig. 2** Location map of the studied area



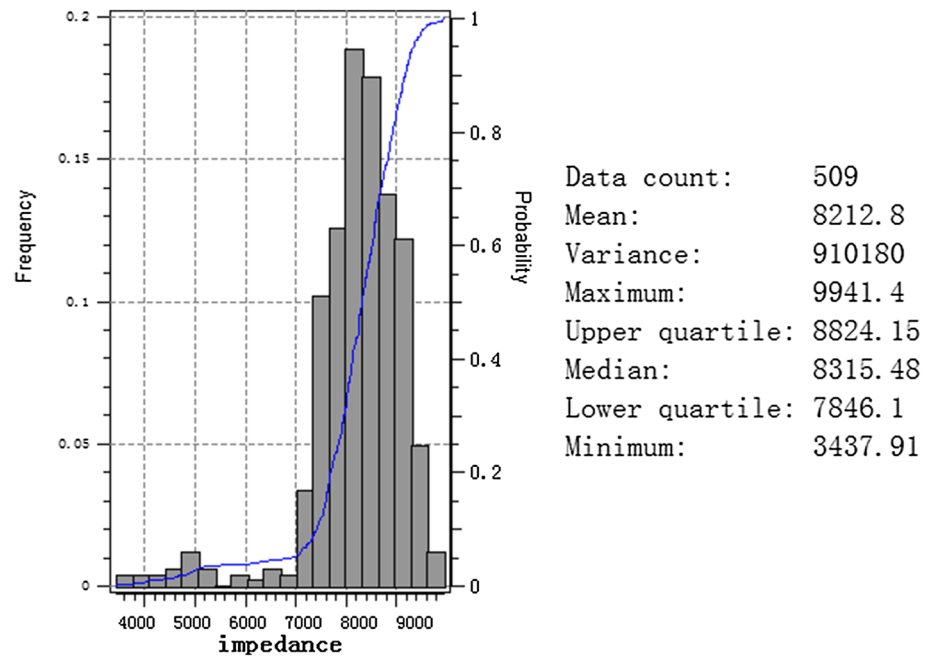
**Fig. 3** Stratigraphic chart of Tahe area



**Fig. 4** 3-D work area with four wells, *x* axis and *y* axis indicate east–west direction and south–north direction, respectively. *z* axis indicates vertical direction



**Fig. 5** Basic descriptive statistics of sonic impedance, *blue curve* is the cumulative distribution function



**Table 1** Emission matrix

	Strong	Weak
Mudstone	222	98
Sandstone	28	131
Conglomerate	5	25

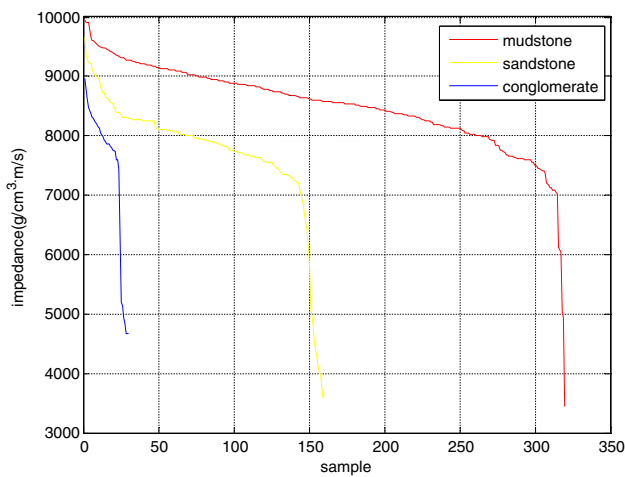
**Table 2** Emission probability

	Strong	Weak
Mudstone	0.6938	0.3063
Sandstone	0.1761	0.8239
Conglomerate	0.1667	0.8333

observed sequence is determined by *B*. As a result, the SHMC can be expressed as  $\lambda = (A, B, C)$ . We consider a first-order neighborhood, which contains six neighbors in 3-D space. Transition matrix *A* can be computed by Eq. (8), where  $l = 6$ . By using the transition probabilities from Fig. 7, the final result of transition matrix *A* with lag  $h = 1, 2, 3, \dots, l$  is

$$A = \begin{pmatrix} 0.7952 & 0.2011 & 0.0037 \\ 0.2534 & 0.6967 & 0.0499 \\ 0.1271 & 0.0608 & 0.8121 \end{pmatrix}$$

where the main diagonal elements indicate the probabilities transfer between same reservoir categorical variables; 0.2011, for example, is the probability transfer from mudstone to sandstone. Emission matrix *B* is



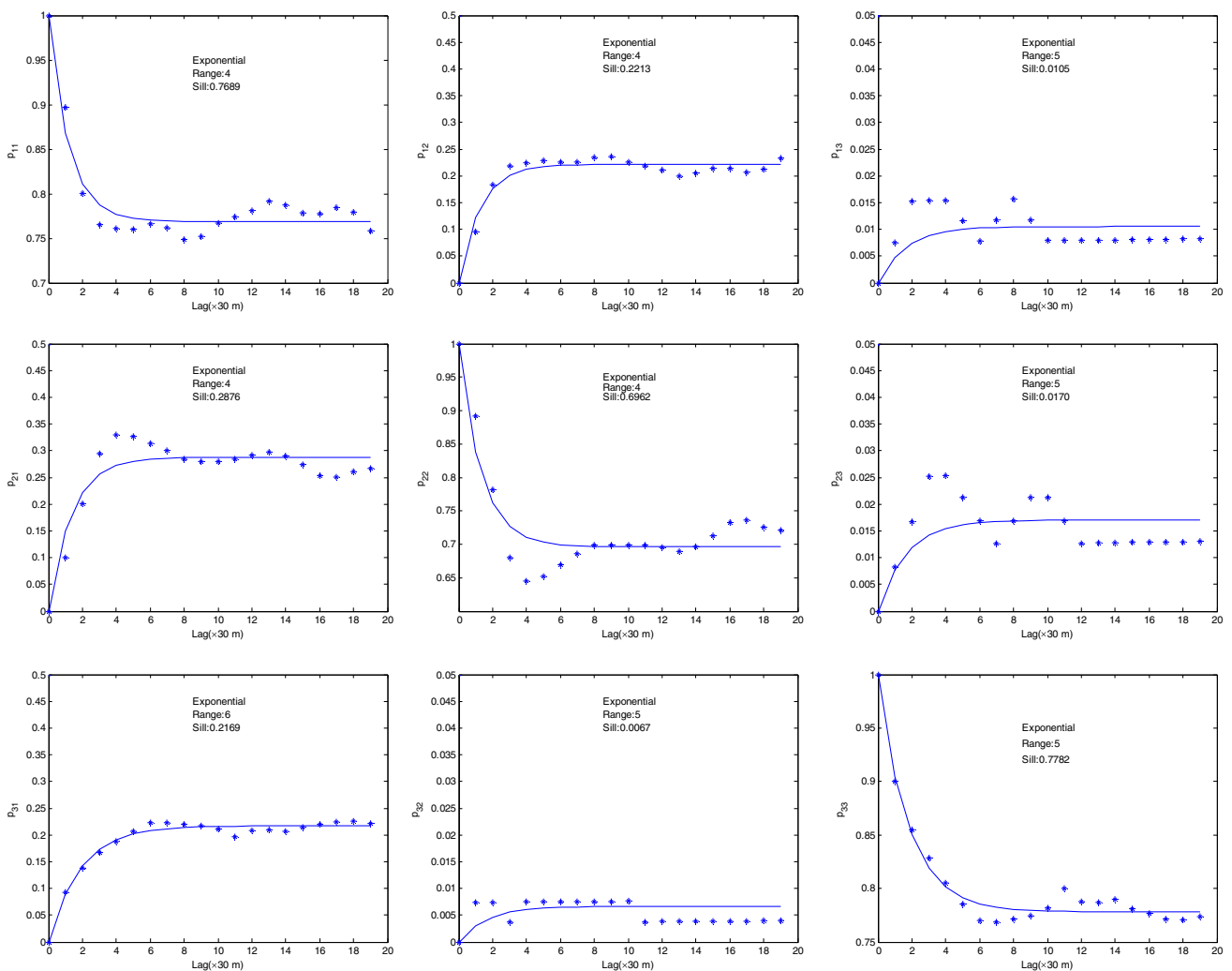
$$B = \begin{pmatrix} 0.6938 & 0.3063 \\ 0.1761 & 0.8239 \\ 0.1667 & 0.8333 \end{pmatrix}$$

which is given in Table 2. Initial probabilities (prior probabilities)  $C$  are

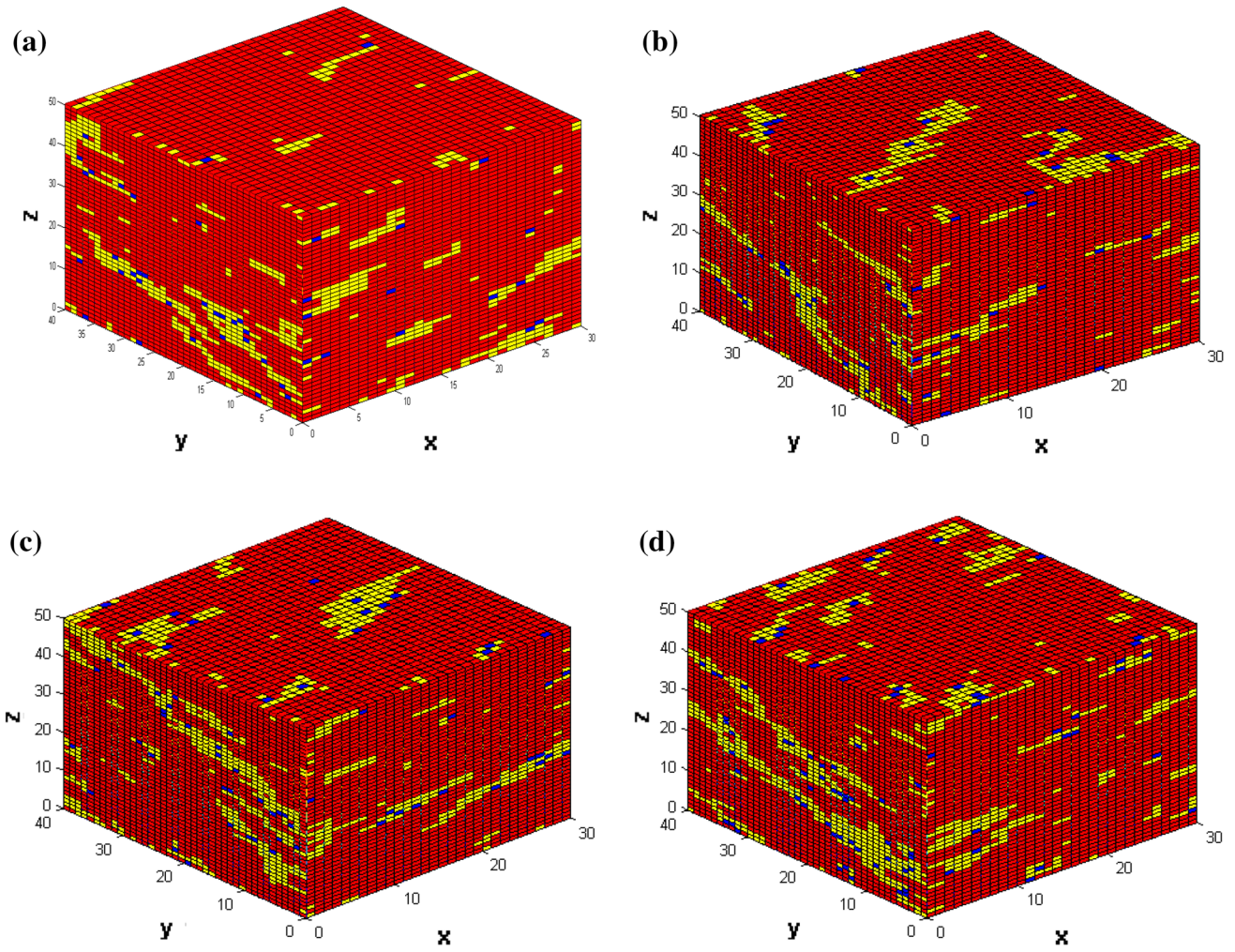
$$C = (0.6287 \quad 0.3124 \quad 0.0589)^T.$$

By using Viterbi algorithm, four realizations are simulated for work area (Fig. 8). In order to compute the simulation accuracy of the SHMC method, another well, S75, is added in the middle of the section between wells 66 and 67 (Fig. 9). Note that “S75” is a short notation for “well 75.” We have got 121 lithofacies samples in this well. The newly obtained log data can be used as validation sets. By comparing the estimated lithofacies in S75

**Fig. 6** Impedance of each reservoir categorical variable. Conglomerate: left, sandstone: middle, mudstone: right

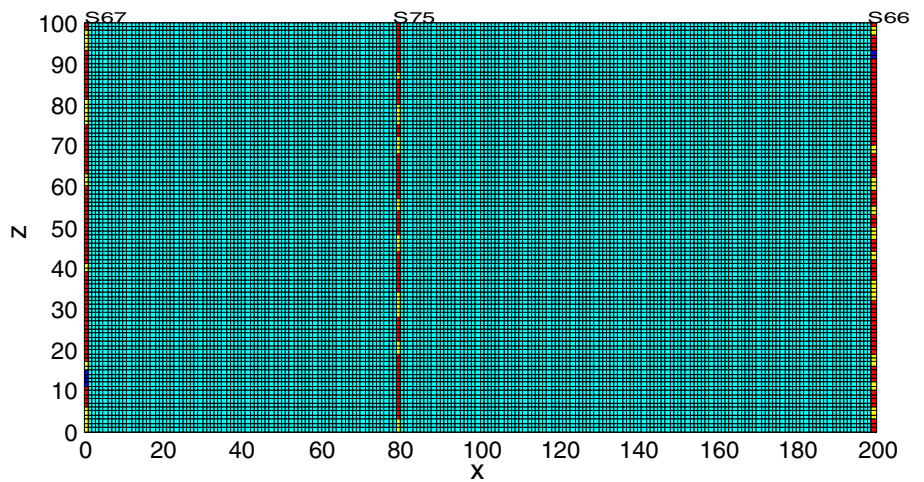


**Fig. 7** Experimental transiagrams and fitted models



**Fig. 8** Four realizations implemented by SHMC. Mudstone: *red*, sandstone: *yellow*, conglomerate: *blue*

**Fig. 9** A section to be simulated across three wells





**Table 3** SHMC classification accuracy

	Simulation (a)	Simulation (b)	Simulation (c)	Simulation (d)	Average
Mudstone	66/85	58/85	61/85	51/85	55/85
Sandstone	18/31	25/31	20/31	17/31	21/31
Conglomerate	2/5	0/5	0/5	1/5	1/5
Overall	88/121	83/121	81/121	69/121	<b>77/121 (63.64 %)</b>

The value corresponding to the overall average classification accuracy is in bold

location with the true facies, we obtain the classification accuracy, which can be defined as

$$\text{Accuracy} = \frac{\# \text{correct classification}}{\# \text{validation sets}}$$

The average prediction accuracy is 63.64 % according to four stochastic simulation results (Table 3).

**Comparison analysis**

To better demonstrate the superiority of the SHMC method, a comparison study has also been conducted. We

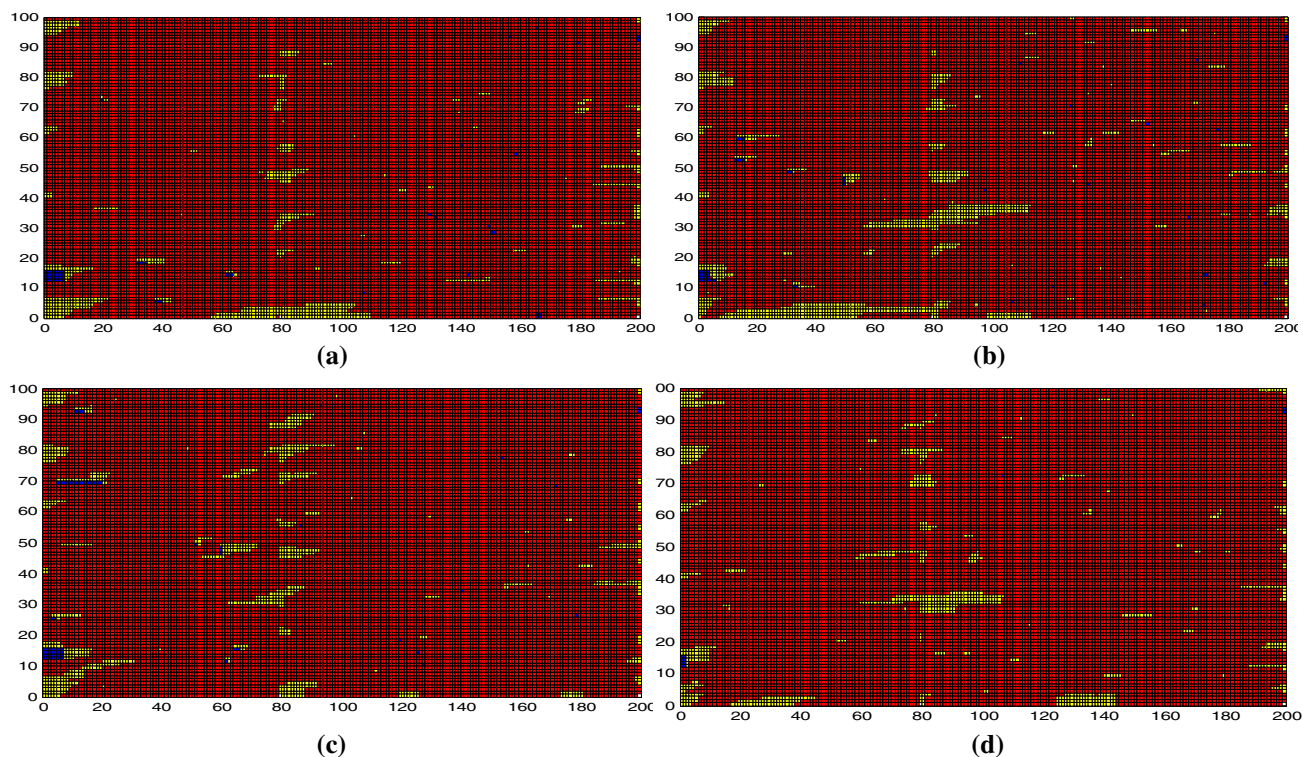
**Table 4** Classification accuracy comparison

	Mudstone	Sandstone	Conglomerate	Overall
<b>SMC</b>				
Simulation (a)	56/85	15/31	2/5	73/121
Simulation (b)	48/85	19/31	0/5	67/121
Simulation (c)	63/85	11/31	3/5	77/121
Simulation (d)	51/85	21/31	0/5	72/121
Average	55/85	17/31	1/5	<b>72/121 (59.50 %)</b>
<b>SHMC</b>				
Simulation (a)	66/85	18/31	1/5	85/121
Simulation (b)	62/85	16/31	3/5	81/121
Simulation (c)	69/85	15/31	0/5	84/121
Simulation (d)	61/85	22/31	2/5	85/121
Average	58/85	19/31	1/5	<b>78/121 (64.46 %)</b>

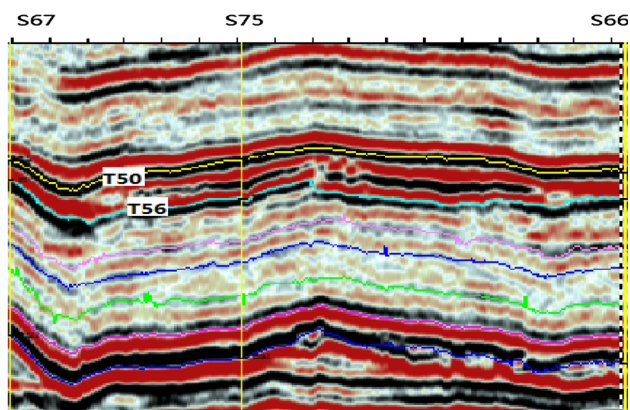
The values corresponding to the overall average classification accuracy are in bold

partition this area into a 100 × 200 grid system, with each unit denoting a 1 m × 4.5 m subsection (Liang 2014). At first, there is no auxiliary information for integration. Thus, we use SMC defined by Eq. (8) for estimation of petroleum reservoir categorical variables. The simulation results obtained by three conditional wells have been shown in Fig. 10. It is obvious that conditional data have played a role in controlling the distribution of lithofacies near the wells. However, the further counterparts are fragmented and random in the grid. The average prediction accuracy is 59.50 % according to four stochastic simulation results (Table 4).

For comparison, the SHMC method has been applied by adding seismic data. Through stratum calibration, time–depth conversion, as well as wave impedance inversion, a seismic section across three wells can be obtained (Fig. 11). Using the impedance partition criterion, we can compute the emission matrix *B*. The entries in this matrix are  $\Pr(w_s|f_s)$ , which can be used in Eq. (7) to calculate the posterior conditional probability combining with Eq. (8). Simulation results have been shown in Fig. 12. The average prediction accuracy increases up to 64.46 % according to four stochastic simulation results (Table 4). Unlike Fig. 10, lithofacies distribution displays certain patterns in random results. More specifically, the distribution of sandstone is not continuous as a whole, with extension about 400–500 m in the horizontal direction. In addition, the section can be divided into three small layers from top to bottom. The middle layer, with the thickness of around 20 m, is twice as thick as the upper and lower ones. Each layer is stacked in space, not connected with each other. As the background lithofacies, mudstone exists widely in this area. Conglomerate, on the other hand, is not well developed due to petroleum geology condition. The simulation results and wave impedance inversion results have good correspondence, which demonstrates that the SHMC model is preferred in the estimation of petroleum reservoir categorical variables.



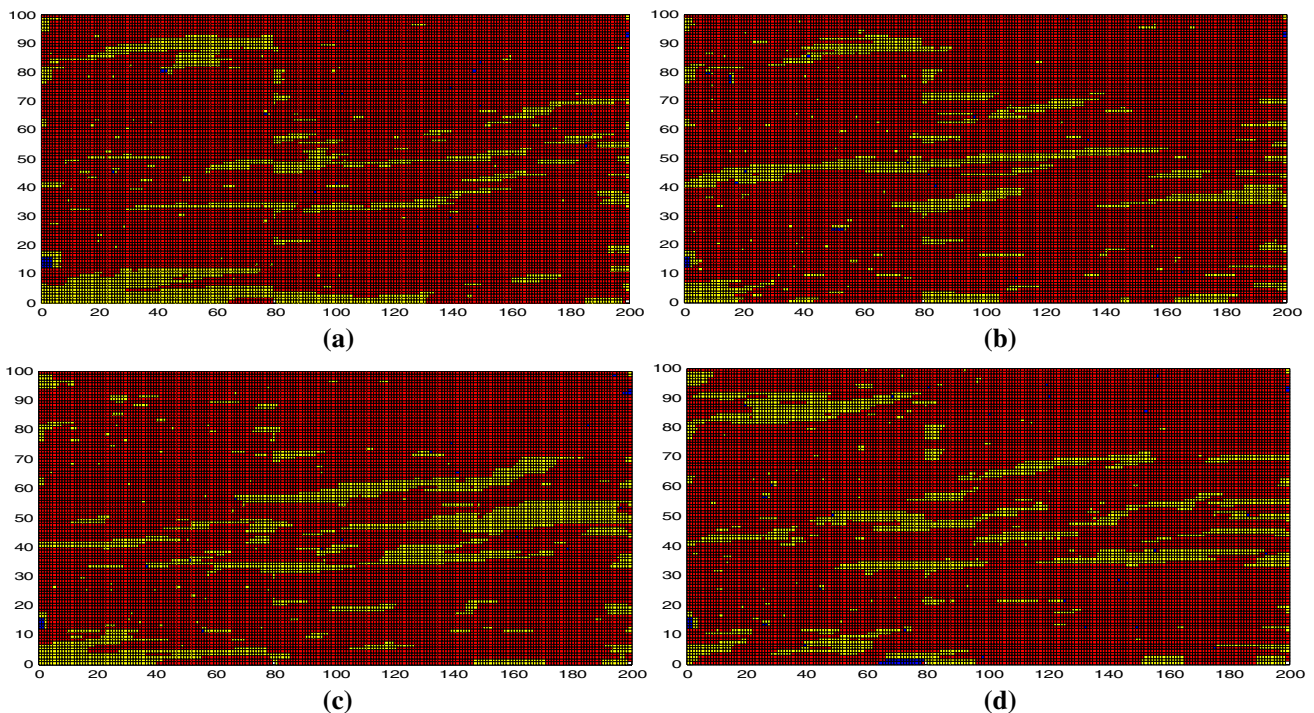
**Fig. 10** Four stochastic simulation results based on three wells



**Fig. 11** A seismic section across three wells

## Conclusions

We have presented an SHMC model for geological facies modeling. This combines spatial Markov chain theory and Bayes estimation. We have adopted the specification of earlier published hidden Markov models. SHMC is based on neighborhood and cliques and has a solid theoretical foundation. Unlike SMC, SHMC integrates well data and geological conceptual data (sonic impedance) by using Viterbi algorithm. In our research, the sonic impedance is divided into two classes: strong and weak, which is regarded as observed variable. Experimental transiograms and fitted models are given according to 509 samples, which are used to compute prior conditional probabilities (transition probabilities). Compared with SMC based on



**Fig. 12** Four stochastic simulation results based on three wells and sonic impedance

well data, the SHMC method performs superiority both in prediction accuracy and reflecting the geological sedimentation process by integrating auxiliary information.

**Acknowledgments** This study is sponsored by the Fundamental Research Funds for the Central Universities of Central South University (No. 2016zzts011) and National Science and Technology Major Project of China (No. 2011ZX05002-005-006). The authors are indebted to Dr. Kan Wu and Dr. Dongdong Chen for their valuable help on transiograms fitting and three-dimensional stochastic simulation. Finally, the authors gratefully thank the editor-in-chief and two anonymous reviewers for their constructive comments and suggestions, which have profoundly improved the composition of this manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Arinkoola AO, Onuh HM, Ogbe DO (2015) Quantifying uncertainty in infill well placement using numerical simulation and experimental design: case study. *J Petrol Explor Prod Technol* 8:1–15
- Blake A, Kohli P, Rother C (2011) Markov random fields for vision and image processing. The MIT Press, Cambridge, pp 11–22
- Carle SF, Fogg GE (1997) Modeling spatial variability with one and multidimensional continuous-lag Markov chains. *Math Geol* 29(7):891–918
- Eidsvik J, Mukerji T, Switzer P (2004) Estimation of geological attributes from a well log: an application of hidden Markov chains. *Math Geol* 36(3):379–397
- Elfeki A, Dekking M (2001) A Markov chain model for subsurface characterization: theory and applications. *Math Geol* 33(5):569–589
- Huang X, Wang Z, Guo J (2016a) Theoretical generalization of Markov chain random field from potential function perspective. *J Cent South Univ* 23(1):189–200
- Huang X, Wang Z, Guo J (2016b) Prediction of categorical spatial data via Bayesian updating. *Int J Geogr Inf Sci* 30(7):1426–1449
- Kolbjørnsen O, Stien M, Kjøsberg H, Fjellvoll B, Abrahamsen P (2014) Using multiple grids in Markov mesh facies modeling. *Math Geosci* 46(2):205–225
- Li W (2006) Transiogram: a spatial relationship measure for categorical data. *Int J Geogr Inf Sci* 20(6):693–699
- Li W (2007) Markov chain random fields for estimation of categorical variables. *Math Geol* 39(3):321–335
- Li J, Xiong L, Fang S, Tang L, Huo H (2010) Lithology stochastic simulation based on Markov chain models integrated with multi-scale data. *Acta Pet Sin* 31(1):73–77 (in Chinese)
- Li J, Yang X, Zhang X, Xiong L (2012) Lithologic stochastic simulation based on the three-dimensional Markov chain model. *Acta Pet Sin* 33(5):846–853 (in Chinese)
- Li W, Zhang C, Willig MR, Dey DK, Wang G, You L (2015) Bayesian Markov chain random field cosimulation for improving land cover classification accuracy. *Math Geosci* 47(2):123–148
- Liang Y (2014) Stochastic simulation of reservoir lithofacies based on the bidirectional Markov chain model. Master's Thesis, Central South University, Changsha, China

- Ma Y, Jones TA (2001) Teacher's aide: modeling hole-effect variograms of lithology-indicator variables. *Math Geol* 33(5):631–648
- Pickard DK (1980) Unilateral Markov fields. *Adv Appl Probab* 12(3):655–671
- Salomão MC, Remacre AZ (2001) The use of discrete Markov random fields in reservoir characterization. *J Petrol Sci Eng* 32(s 2–4):257–264
- Stien M, Kolbjørnsen O (2011) Facies modeling using a Markov mesh model specification. *Math Geosci* 43(43):611–624
- Tjelmeland H, Besag J (1998) Markov random fields with higher-order interactions. *Scand J Stat* 25(25):415–433
- Weissmann GS, Fogg GE (1999) Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *J Hydrol* 226(1):48–65