A Markov Chain Model for Subsurface Characterization: Theory and Applications[1]

Amro Elfeki[2] and Michel Dekking[3]

This paper proposes an extension of a single coupled Markov chain model to characterize heterogeneity of geological formations, and to make conditioning on any number of well data possible. The methodology is based on the concept of conditioning a Markov chain on the future states. Because the conditioning is performed in an explicit way, the methodology is efficient in terms of computer time and storage. Applications to synthetic and field data show good results.

INTRODUCTION

At present, a variety of techniques are available to characterize reservoir heterogeneity. Some selected ones can be found in the literature of hydrogeology (see Neuman, 1980) and reservoir engineering (see Haldorsen, and Damsleth, 1990, and Deutsch and Journel, 1992). Most of these techniques rely on the use of variogram or autocovariance function to describe the spatial structure of reservoir heterogeneity. An alternative to describe the spatial structure is by the use of Markov chains. Markov chains are applied in geology to model discrete variables such as lithologies or facies. The Markov chain model does not use variograms or autocovariance functions to quantify the spatial structures as most of the available models do. Instead, it uses conditional probabilities. Conditional probabilities have the advantage that they are interpreted geologically much easier than variogram or

---

[2]Current address: Department of Hydrology and Ecology, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P. O. Box 5028, 2600GA Delft, The Netherlands. On leave from Faculty of Engineering, Mansoura University, Egypt. e-mail: aelfeki@mum.mans.eun.eg
[3]Thomas Stieltjes Institute for Mathematics and Faculty of ITS, Department of Probability, Statistics and Operation Research, Delft University of Technology, P. O. Box 5031, 2600GA Delft, The Netherlands. e-mail: f.m.dekking@its.tudelft.nl

autocovariance functions. This is the reason for their popularity in the geological community.

This paper proposes an extension of the coupled Markov chain model, developed by Elfeki (1996) to characterize heterogeneity of natural formations, to conditioning on any number of well data. The coupled Markov chain model is also an extension of the one-dimensional Markov chain model used by Krumbein (1967) to synthesize a stratigraphic sequence. The methodology is based on a Markov chain that is conditioned on the future states. The conditioning is performed in an explicit way that makes it efficient in terms of computer time and storage. In the next sections, the basic concepts of the classical one-dimensional Markov chain and the coupled Markov chains theories are presented followed by the concept of conditioning of a Markov chain on the future states. Some applications on a hypothetical case study and on real outcrop data are presented. For other work on Markov chains to model geological formations see Lin and Harbaugh (1984) and Moss (1990). Recent directions regarding Markov chains applications in geology can be found in Carle (1996) and Carle and Fogg (1996).

ONE-DIMENSIONAL MARKOV CHAINS

A Markov chain is a probabilistic model that exhibits a special type of dependence: given the present the future does not dependent on the past. In formulas, let $Z_0, Z_1, Z_2, \ldots, Z_m$ be a sequence of random variables taking values in the state space $\{S_1, S_2, \ldots, S_n\}$. The sequence is a Markov chain or Markov process if

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_{i-2} = S_n, Z_{i-3} = S_r, \ldots, Z_0 = S_p)$$
$$= \Pr(Z_i = S_k \mid Z_{i-1} = S_l) =: p_{lk} \qquad (1)$$

where the symbol "|" is the symbol for conditional probability.

Transition Matrix and Stationary Probabilities

In one-dimensional problems a Markov chain is described by a single transition probability matrix. Transition probabilities correspond to relative frequencies of transitions from a certain state to another state. These transition probabilities can be arranged in a square matrix form,

$$\boldsymbol{p} = \begin{bmatrix} p_{11} & p_{12} & \cdot & \cdot & p_{1n} \\ p_{21} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & p_{lk} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{n1} & \cdot & \cdot & \cdot & p_{nn} \end{bmatrix}$$

where p_{lk} denotes the probability of transition from state S_l to state S_k, and n is the number of states in the system. Thus the probability of a transition from S_1 to $S_1, S_2, \ldots, S_n$ is given by $p_{1l}, l = 1, 2, \ldots, n$ in the first row and so on. The matrix p has to fulfil specific properties: (1) its elements are non-negative, $p_{lk} \geq 0$; (2) the elements of each row sum up to one or

$$\sum_{k=1}^{n} p_{lk} = 1 \tag{2}$$

The transition probabilities considered in the previous section are called single step transitions. One considers also N-step transitions, which means that transitions from a state to another take place in N steps. The N-step transition probabilities can be obtained by multiplying the single-step transition probability matrix by itself N times. Under some mild conditions on the transition matrix (aperiodicity and irreducibility), the successive multiplication leads to identical rows $(w_1, w_2, \ldots, w_n)$. So the $w_k (k = 1, 2, \ldots, n)$ are given by

$$\lim_{N \to \infty} p_{lk}^{(N)} = w_k \tag{3}$$

and are called stationary probabilities. The w_k are no longer dependent on the initial state S_l. The stationary probabilities may also be found by solving the equations

$$\sum_{l=1}^{n} w_l p_{lk} = w_k, \quad k = 1, \ldots, n \tag{4}$$

subject to the conditions $w_k \geq 0$ and $\sum_k w_k = 1$.

One-Dimensional Markov Chain Conditioned on Future States

Consider a one-dimensional series of events that are Markovian (Fig. 1). The probability of cell i to be in state S_k, given that the previous cell $i - 1$ is in state S_l and cell N is in state S_q can be expressed mathematically as
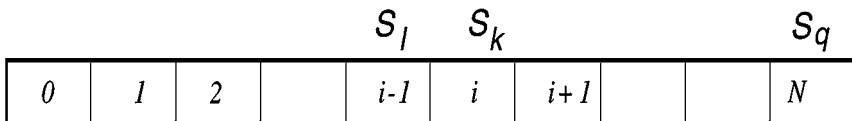
$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q)$$

				S_l	S_k				S_q
0	1	2		i-1	i	i+1			N

Figure 1. Numbering series of events for a one-dimensional Markov chain.

This probability can be written in terms of joint probabilities as

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q) = \frac{\Pr(Z_{i-1} = S_l, Z_i = S_k, Z_N = S_q)}{\Pr(Z_{i-1} = S_l, Z_N = S_q)} \quad (5)$$

One can factorize the numerator of Equation (5) leading to

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q)$$
$$= \frac{\Pr(Z_N = S_q \mid Z_{i-1} = S_l, Z_i = S_k)\Pr(Z_{i-1} = S_l, Z_i = S_k)}{\Pr(Z_{i-1} = S_l, Z_N = S_q)} \quad (6)$$

By applying the Markovian property on the conditional probability in the numerator of Equation (6), one obtains

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q)$$
$$= \frac{\Pr(Z_N = S_q \mid Z_i = S_k) \cdot \Pr(Z_{i-1} = S_l, Z_i = S_k)}{\Pr(Z_{i-1} = S_l, Z_N = S_q)} \quad (7)$$

The joint distributions in the numerator and the denominator can be expressed in terms of conditional probabilities and marginal probabilities as

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q)$$
$$= \frac{\Pr(Z_N = S_q \mid Z_i = S_k) \cdot \Pr(Z_i = S_k \mid Z_{i-1} = S_l) \cdot \Pr(Z_{i-1} = S_l)}{\Pr(Z_N = S_q \mid Z_{i-1} = S_l) \cdot \Pr(Z_{i-1} = S_l)} \quad (8)$$

The conditional probabilities in Equation (8) can be expressed in terms of the matrix p as

$$\Pr(Z_i = S_k \mid Z_{i-1} = S_l, Z_N = S_q) = \frac{p_{kq}^{(N-i)} p_{lk}}{p_{lq}^{(N-i+1)}} \quad (9)$$

where $p_{kq}^{(N-i)}$ is the $(N-i)$-step transition probability, $p_{lk}^{(N-i+1)}$ is the $(N-i+1)$-step transition probability. Equation (9) can be rearranged in the form

$$p_{lk|q} = \frac{p_{lk} p_{kq}^{(N-i)}}{p_{lq}^{(N-i+1)}} \quad (10)$$

where $p_{lk|q}$ is our target, the probability of cell i to be in state S_k, given that the previous cell $i-1$ is in state S_l and the future cell N is in state S_q.

In Equation (10) when cell N is far from cell i the terms $p_{lq}^{(N-i+1)}$ and $p_{kq}^{(N-i)}$ cancel because they are almost equal to the stationary probability w_q. However, when we get closer to cell N, its state starts to play a role and simulation results are influenced by the state at that cell.

COUPLED MARKOV CHAIN THEORY

The coupled chain describes the joint behavior of a pair of independent systems, each evolving according to the laws of a classical Markov chain (Billingsley, 1995). Consider two Markov chains $(X_i),(Y_j)$ both defined on the state space $\{S_1, S_2, \ldots, S_n\}$ and having the positive transition probability defined as

$$\Pr(X_{i+1} = S_k, Y_{j+1} = S_f \mid X_i = S_l, Y_j = S_m) = p_{lm,kf} \qquad (11)$$

The coupled transition probability $p_{lm,kf}$ on the state space $\{S_1, S_2, \ldots, S_n\} \times \{S_1, S_2, \ldots, S_n\}$ is given by

$$p_{lm,kf} = p_{lk} \cdot p_{mf} \qquad (12)$$

These transition probabilities from a stochastic matrix.

Coupled Markov Chain on a Lattice

Two-coupled one-dimensional Markov chains (X_i) and (Y_j) can be used to construct a two dimensional spatial stochastic process on a lattice $(Z_{i,j})$. Consider a two-dimensional domain of cells as shown in Figure 2. Each cell has a row number j and a column number i. Consider also a given number of geological materials, say n. Geological materials are coded as numbers. The word state, in this text, describes a certain geological unit, lithology, or bedding type. The (X_i) chain describes the structure of the geological unit in the horizontal direction. We write

$$p_{lk}^h = \Pr(X_{i+1} = S_k \mid X_i = S_l) \qquad (13)$$

Similarly, the (Y_i) chain describes the structure in the vertical direction and we write

$$p_{mk}^v = \Pr(Y_{j+1} = S_k \mid Y_j = S_m) \qquad (14)$$

The stochastic process (Z_{ij}) is obtained by coupling the (X_i) and (Y_j) chain, but forcing these chains to move to equal states. Hence,

$$\Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m)$$
$$= C \Pr(X_i = S_k \mid X_{i-1} = S_l) \Pr(Y_j = S_k \mid Y_{j-1} = S_m) \qquad (15)$$

1,1									Nx,1
				i,j-1					
			i-1,j	i,j					Nx,j
1,Ny									Nx,Ny

Coupled Markov Chain without Conditioning on the Right Boundary.

1,1									Nx,1
				i,j-1					
			i-1,j	i,j					Nx,j
1,Ny									Nx,Ny

Coupled Markov Chain with Conditioning on the Right Boundary.

Figure 2. Numbering system in two-dimensional domain for the coupled Markov chain. Unconditional Markov chain (top) and conditional Markov chain on the states of the future (bottom). Dark grey cells are known boundary cells, light grey cells are known cells inside the domain (previously generated, the past), white cells are unknown cells. The future state used to determine the state of cell (i,j) is cell (N_x,j).

where C is a normalizing constant that arises by not permitting transitions in the (X_i) and (Y_j) chain to different states. Hence,

$$C = \left(\sum_{f=1}^{n} p_{lf}^{h} \cdot p_{mf}^{v} \right)^{-1} \tag{16}$$

Combining Equation (15) and Equation (16) the required probability is therefore

$$p_{lm,k} := \Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m)$$

$$= \frac{p_{lk}^{h} \cdot p_{mk}^{v}}{\sum_f p_{lf}^{h} \cdot p_{mf}^{v}} \quad k = 1, \ldots, n \tag{17}$$

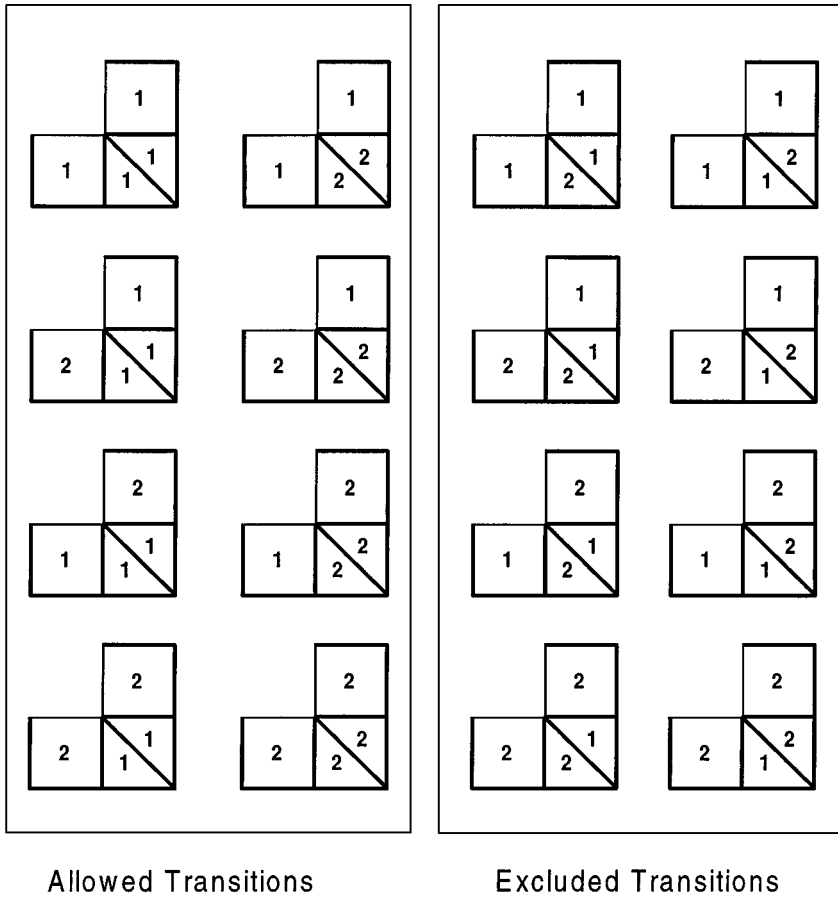## Allowed Transitions   Excluded Transitions

Figure 3. Transition scheme in the two-state model.

As an example, suppose we have two chains that are operating on two different states coded as 1 and 2. The possible states at cell (i, j) given that the state of the previous cell $(i-1, j)$ was 1, and top cell $(i, j-1)$ was in state 1 as well, are displayed in Figure 3 (top row). It is possible that one obtains the same state 1 or 2 from both chains as in the first two possibilities in Figure 3 (top row), or one could get different states from the chains as in the last two possibilities in the top row. Here, for example, one has

$$p_{12,1} = \Pr(Z_{i,j} = 1 \mid Z_{i-1,j} = 1, Z_{i,j-1} = 2) = \frac{p_{11}^h \cdot p_{21}^v}{p_{11}^h \cdot p_{21}^v + p_{12}^h \cdot p_{22}^v} \quad (18)$$

Conditioning the Coupled Markov Chain on Two Wells

Since it is already not evident how to define a notion for "future" in two dimensions, it is not straightforward to extend the conditioning of one-dimensional Markov chains on future states to the coupled chain. We shall therefore mainly consider a very simple and computationally cheap approximative way by performing conditioning of the horizontal chain first, and coupling the conditioned horizontal chain with the vertical chain afterward. The expression of the conditional probability in the coupled chain, according to this way, given the cell on the right hand side boundary is given by

$$\Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m, Z_{N_x,j} = S_q) = C'\Pr(Z_{i,j}$$

$$= S_k \mid Z_{i-1,j} = S_l, Z_{N_x,j} = S_q) \cdot \Pr(Z_{i,j} = S_k \mid Z_{i,j-1} = S_m) \quad (19)$$

Here C' is a normalising constant as in Equation (16).

By inserting the expression derived in Equation (10) into Equation (19) and proceeding as in the previous section, we obtain

$$\Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m, Z_{N_x,j} = S_q) = C' \frac{p_{lk}^h \cdot p_{kq}^{h(N_x-i)}}{p_{kq}^{h(N_x-i+1)}} \cdot p_{mk}^v$$

$$(20)$$

Computing C' as in the previous section and canceling the numerators, we finally find

$$p_{lm,k|q} := \Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m, Z_{N_x,j} = S_q)$$

$$= \frac{p_{lk}^h \cdot p_{kq}^{h(N_x-i)} \cdot p_{mk}^v}{\sum_f p_{lf}^h \cdot p_{fq}^{h(N_x-i)} \cdot p_{mf}^v} \quad k = 1, \ldots, n \quad (21)$$

For exact conditioning it is useful to note that the coupled Markov chain (Z_{ij}) is an example of unilateral Markov field (Pickard, 1980). Such Markov fields can also be described by a (one-dimensional) Markov chain in a random environment (this observation has also been made in Galbraith and Walley, 1976). For each state $S_m, m = 1, \ldots, n$ define a transition matrix p^m by

$$p_{lk}^m := p_{lm,k} = \Pr(z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m) \quad (22)$$

[cf. Equation (17)]. With this point of view it is possible to compute an N-step transition probability for the process for a fixed j given the "past"—cell $(i-1,j)$

and the previous row—by

$$\Pr\big(Z_{i+N,j} = S_k \mid Z_{i,j} = S_l, Z_{i+1,j-1} = S_{m(1)}, \ldots, Z_{i+N,j-1} = S_{m(N)}\big)$$

$$= \big(p_{lk}^{m(1)} p_{lk}^{m(2)} \cdots p_{lk}^{m(N)}\big)_{lk} \tag{23}$$

where $p^{m(1)}, p^{m(2)} \cdots p^{m(N)}$ is the ordinary matrix product of $p^{m(1)}, p^{m(2)}, \cdots, p^{m(N)}$. Equation (23) follows by induction from the case $N = 2$, which is proved with manupulations similar to those to derive Equations (5)–(7). Now we can just as in the one-dimensional case condition on the "future," defining the future of cell (i, j) to be cell (N_x, j) (cf. Fig. 2). We obtain that

$$\Pr\big(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_{m(1)}, \ldots, Z_{N_x,j-1}$$

$$= S_{m(N_x-i+1)}, Z_{N_x,j} = S_q\big) = \frac{p_{lk}^{m(1)} \cdot \big(p^{m(2)} \cdots p^{m(N_x-i+1)}\big)_{kq}}{\big(p^{m(1)} \cdots p^{m(N_x-i+1)}\big)_{lq}} \tag{24}$$

It is clear that this exact conditioning will be computationally more expensive.

INFERENCE OF STATISTICAL PARAMETERS FROM A GEOLOGICAL SYSTEM

This section is similar to the description given by Elfeki (1996) for estimation of model parameters from field observations. For the sake of completeness, parameter estimation is explained once again in the present context. A Markov chain is described completely when the state space, transition probabilities, and initial probabilities are given. The initial probabilities will be chosen equal to the stationary probabilities. We will not need to estimate an initial distribution as we will do simulations conditioned on well log and surface data. For a geological system represented by a Markov chain, one has to perform the following steps. First, the set of possible states of the system $\{S_1, S_2, \ldots, S_n\}$ must be defined. Second, the probability p_{lk} of going from a state S_l to state S_k in one interval must be estimated. Finally, the stationary probabilities w_k are determined either by estimation from the data or by calculation from the transition probabilities. In practical applications, transition and stationary probabilities of a geological system can be estimated from well logs, bore holes, surface and subsurface mapping, or from geological images synthesized by information derived from geologically similar sites or analogous outcrops. The estimation procedure is given below.

Estimation of Transition Probabilities

The vertical transition probability matrix can be estimated from well logs. The tally matrix of vertical transitions is obtained by superimposing a vertical

line with equidistant points along the well with a chosen sampling interval. The transition frequencies between the states are calculated by counting how many times a given state S_l is followed by itself or the other states S_k in the system, and then divided by the total number of transitions,

$$p_{lk}^v = \frac{T_{lk}^v}{\sum_{q=1}^u T_{lq}^v} \tag{25}$$

where T_{lk}^v is the number of observed transitions from S_l to S_k in the vertical direction.

Similarly, the horizontal transition probability matrix can be estimated from geological maps. Maps that show formation extensions in the horizontal plane may be obtained from geological surveys. On the map plan, a transect is defined where the subsurface profile is required. On the transect, a similar procedure is performed as in the case of vertical transitions. A horizontal line with equidistant points at a proper sampling interval for the horizontal direction is superimposed over the map. The transitions between different states are counted and the horizontal transition probability matrix is computed using Equation (25) with superscript h instead of v.

Estimation of the Sampling Intervals

Estimation of the proper sampling intervals is a trial and error procedure. There is no specific rule. Perhaps the proper sampling interval in the vertical direction would be less than or equal to the minimum thickness of the geologic unit found in the well log in order to be reasonably reproduced in the simulation. Similarly, the proper sampling interval in the horizontal direction would be less than or equal to the minimum length of a geological unit found on a planar geological map.

THE ALGORITHM

A procedure for Monte Carlo sampling to implement this methodology is presented. Refer to Figure 2 during the description of the algorithm. The procedure for conditional simulation on two neighboring wells is as follows:

Step 1: The two-dimensional domain is discretized using proper sampling intervals.

Step 2: Well data is inserted in their locations at the well on the left side of the domain at $(1, j)$, $j = 2, \ldots, N_y$, the well at the right side of the domain at (N_x, j), $j = 2, \ldots, N_y$, and the top surface information at location $(i, 1)$, $i = 1, \ldots, N_x$.

Step 3: Generate the rest of the cells inside the domain that is numbered (i, j), $i = 2, \ldots, N_x - 1$ and $j = 2, \ldots, N_y$ rowwise using the

conditional distribution $\Pr(Z_{i,j} = S_k \mid Z_{i-1,j} = S_l, Z_{i,j-1} = S_m,$ $Z_{N_x,j} = S_q)$, given that the states at $(i-1, j)$, $(i, j-1)$, and (N_x, j) are known. The four-index conditional probability $p_{lm,k\mid q}$ is calculated by Equation (21). From state S_l at the horizontal neighboring cell $(i-1, j)$, S_m at the vertical neighboring cell $(i, j-1)$ and the state S_q at the cell on the right-hand side boundary (N_x, j), one can determine the succeeding state S_k at cell (i, j). We simulate a state for S_k according to the distribution given by $(p_{lm,r\mid q}: r = 1, \ldots, n)$.

Step 4: The procedure stops after having visited all the cells in the domain between the two given wells at $i = 1$ and $i = N_x$.

Step 5: The same procedure is followed for the next two wells and so on until the domain is filled with the states.
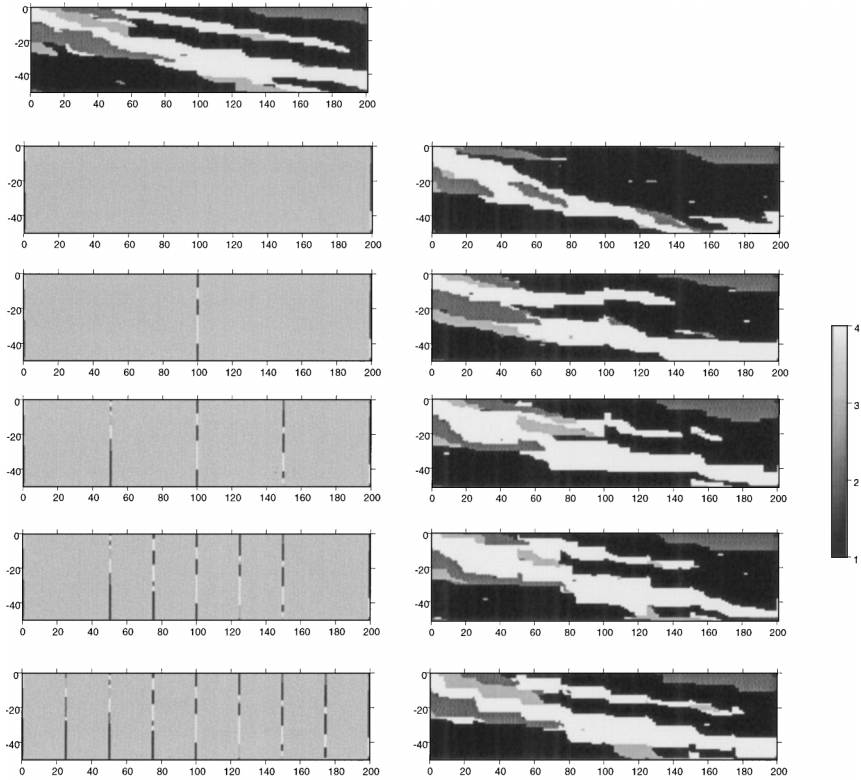
HYPOTHETICAL EXAMPLE

In this section a hypothetical example is presented. In this example, the data needed for the simulation is given in Table 1. A geological cross-section 200 km long and 50 m in depth is considered. The geological system contains four different geological materials coded by 1–4. The transition probabilities in both the horizontal and the vertical directions are displayed in Table 1. These probabilities are used to generate the artificial geological structure. In Table 1 the sampling intervals over which these transitions are applied are given. The artificial geological structure generated by data in Table 1 is shown in Figure 4 (top left).

Figure 4 shows the well locations on the left-hand side. The corresponding stochastically generated realizations using these wells are displayed on the right-hand side. It is clear that by increasing the number of wells, the simulation results improve and become closer to the original image ("real" formation). Figure 4 shows only single realizations conditioned on several wells. In order to evaluate the uncertainty and degree of variability between the realizations, a Monte Carlo approach

**Table 1.** Input Data for the Hypothetical Example

Length of the given section (km) = 200	Depth of the given section (m) = 50
Sampling interval in X-axis (km) = 1	Sampling interval in Y-axis (m) = 1

No. of states = 4

Horizontal transition probability matrix					Vertical transition probability matrix			
State	1	2	3	4	State	1	2	3	4
1	0.980	0.005	0.005	0.010	1	0.900	0.030	0.030	0.040
2	0.010	0.970	0.010	0.010	2	0.040	0.900	0.030	0.030
3	0.020	0.010	0.960	0.010	3	0.030	0.030	0.900	0.040
4	0.010	0.010	0.010	0.970	4	0.040	0.030	0.040	0.900

Figure 4. Hypothetical example (with four different lithologies) showing how the method works when conditioning on more than two wells is performed. Top left is the "real" reservoir and all the next rows show the well locations and lithologies observed at each well (on the left of the figure). The corresponding stochastic simulations (single realizations) conditioned on these wells are shown on the right hand side.

is followed in which one hundred realizations are generated and the ensemble average is calculated over the indicator function of each state. The indicator function of state S_k is given by $I_k(\mathbf{x}) = 1$ if state S_k is located at $\mathbf{x}$ and $I_k(\mathbf{x}) = 0$ otherwise.

The ensemble average over 100 realizations of the indicator function can be considered as a measure of the probability of occurrence of a specific lithology located at specific point in space. When the ensemble average of the indicator function equals one, it is 100% sure that the lithology S_k is located at $\mathbf{x}$, and when the indicator function is zero it is 100% sure that the lithology S_k is not located at $\mathbf{x}$. Figure 5 shows images of the ensemble average of the indicator functions of the four lithologies conditioned on five wells. For the sake of comparison, a single realization is displayed in Figure 5 with the ensemble indicator function of
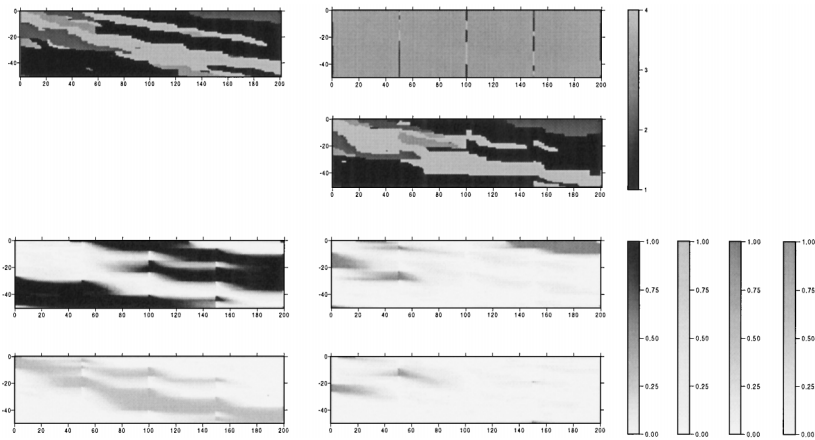
Figure 5. Ensemble averaging over 100 realizations of the hypothetical example shown in Figure 4 (top left) is the "real" reservoir, top right is the well locations, the second row is a single realization, the third rows to the bottom are the ensemble averages of the indicators of each lithology. The grey scale ranges from 0 to 1.

each lithology. It is clear that there are no significant differences between the single realization and the ensemble average of each lithology. This leads to the conclusion that the realizations are not varying so much between one another and so the same pattern is preserved over all the realizations. There are of course some slight variations at the boundaries of the lithology that appeared clearly in the lithology 1 (black) where one can notice fuzzy boundaries that turn to white gradually.

OUTCROP CASE STUDIES

Distal Fluvial Fan Deposits in the Loranca Basin, Spain

The outcrop description can be found in Gozalo and Martinius (1993). A brief description is given below. The outcrop (Fig. 6, top left) shows several sandstone genetic types: fluvial channel deposits, sheet deposits (crevasse-channel and crevasse-splay deposits), and deltaic deposits. One recognizes in the channel deposits two main types: meander-loop deposits and channel-fill deposits. The sheet deposits are sedimentary structures with relatively large extent ($30 \sim 150$ m) and thickness between $0.25 \sim 2$ m. The sheet deposits consists of ripple-laminated medium-grained sand to silt. In the deltic deposits, one finds several stacked sand beds separated by thin mudstone layers ($1 \sim 10$ cm).

The presented methodology is used to generate realizations of two-dimensional (2D) cross-sections of the outcrop. The statistical parameters given in Table 2 are estimated from the schematic outcrop picture displayed in Figure 6
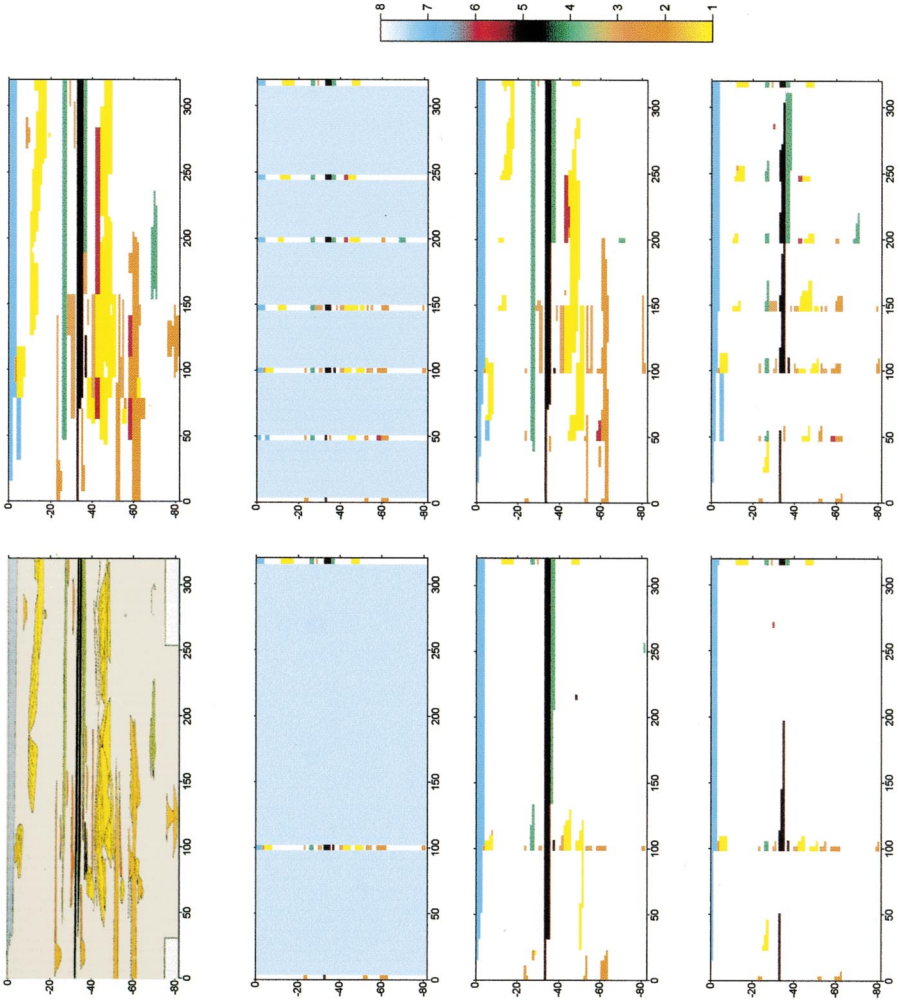
Figure 6. Stochastic simulation (single realizations) of the 2D cross-sectional panel of distal deposits fluvial fan, based on outcrop data (Gozalo and Martinius, 1993). Top left is the observed section, top right is the schematic

Table 2. Input Data Used for the Geological Simulation of the Outcrop (Distal Fluvial Fan Deposits) in the Loranca Basin, Spain[a]

Length of the given section (m) = 320				Depth of the given section (m) = 82			
Sampling interval in X-axis (m) = 4				Sampling interval in Y-axis (m) = 1			
			No. of states = 8				

Horizontal transition probability matrix

State	1	2	3	4	5	6	7	8
1	0.952	0.003	0.003	0.000	0.000	0.003	0.000	0.039
2	0.005	0.923	0.003	0.000	0.000	0.005	0.003	0.061
3	0.000	0.000	0.933	0.010	0.000	0.000	0.000	0.057
4	0.000	0.000	0.000	0.988	0.000	0.000	0.000	0.012
5	0.000	0.000	0.000	0.000	0.995	0.000	0.000	0.005
6	0.018	0.018	0.000	0.000	0.000	0.929	0.000	0.036
7	0.007	0.003	0.000	0.000	0.000	0.000	0.990	0.000
8	0.008	0.004	0.002	0.001	0.001	0.001	0.001	0.982

Vertical transition probability matrix

State	1	2	3	4	5	6	7	8
1	0.783	0.000	0.006	0.000	0.000	0.006	0.000	0.205
2	0.024	0.625	0.022	0.000	0.000	0.018	0.000	0.329
3	0.082	0.000	0.392	0.000	0.000	0.021	0.000	0.505
4	0.000	0.000	0.031	0.510	0.000	0.000	0.000	0.459
5	0.000	0.000	0.038	0.155	0.582	0.000	0.000	0.225
6	0.313	0.143	0.000	0.000	0.000	0.500	0.000	0.045
7	0.017	0.017	0.000	0.000	0.000	0.000	0.698	0.268
8	0.019	0.028	0.020	0.020	0.020	0.011	0.003	0.880

[a]The states in the table are identified in Figure 6.

(top right). Equation (24) is considered for the estimation of these parameters. The simulation of the outcrop has been performed conditioned on three (Fig. 6, second row to the left) and seven (Fig. 6, second row to the right) wells, respectively. The simulation results presented in Figure 6 show good agreements in terms of reproducing the geological features that are present in the outcrop, particularly in

Figure 6. (Continued) section used for simulation purposes. The second row shows artificial well locations (the left image shows three wells and the right image shows seven wells). The third row shows single stochastic realizations conditioned on three wells (left) and seven wells (right) respectivley. The bottom row shows stochastic realizations generated by the old unconditional coupled Markov chain model (Elfeki, 1996). The color scale represents the following: (1) meander-loop deposits, (2) channel-fill deposit, (3) crevasse-channel-splay deposit, (4) lacustrine-deltaic deposit, (5) lacustrine limestone, (6) carbonate palaeosol, (7) gypsum, and (8) mudstone.
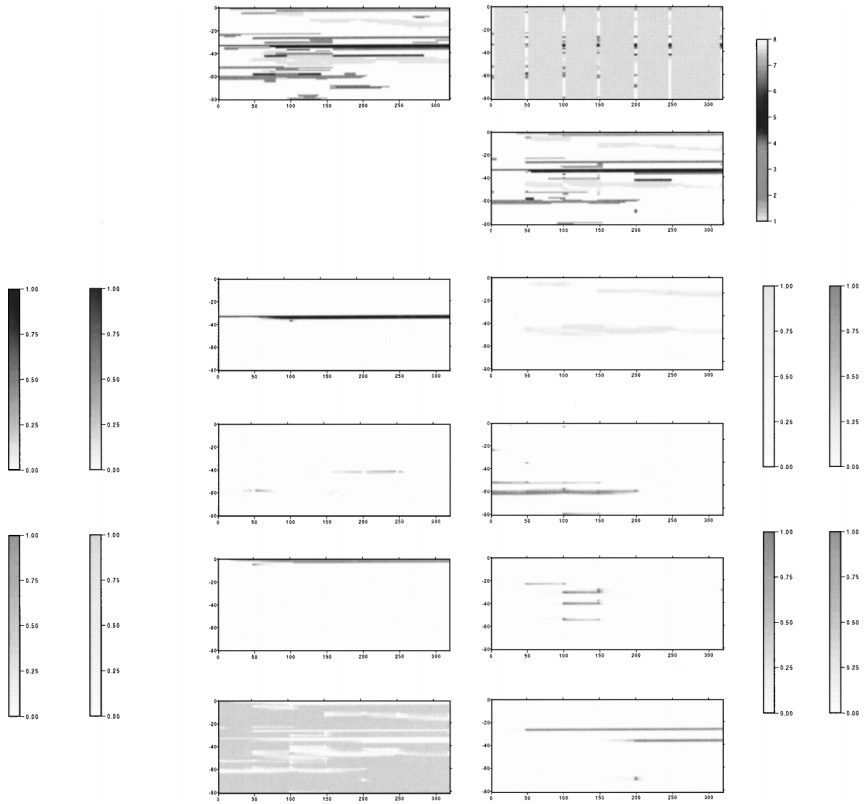
Figure 7. Ensemble averaging over 100 realizations on the 2D cross-sectional panel of distal deposits fluvial fan, based on outcrop data shown in Figure 7. Top left is the schematic outcrop, top right is the well locations, the second row is a single realization, from the third rows to the bottom are the ensemble averages of the indicators of each lithology. The grey scale ranges from 0 to 1.

the case of seven well data (Fig. 6, the third row, the right image), where the well spacing is 50 m. The simulation with three wells (Fig. 6, third row, left image) shows relatively fair agreement for the geological features with long extensions (see the black and the green colours in the simulations). The ensemble average

Figure 8. (Opposite) Stochastic simulation of the two-dimensional cross-sectional panel of the fluvial succession of the medial area of the Tórtola fluvial system, Spain (Martinius, 1996). Top image is the schematization of the real outcrop; the second and the third images are the five well data set and the corresponding simulation (single realization) respectively; the fourth and the fifth images are the eleven wells data set and the corresponding simulation (single realization) respectively. The legend: (1) nonchannelized sheet sandstone bodies, (2) giant-bar sandstone bodies, (3) multistory conglomerate-rich bodies, (4) composite point-bar sandstone bodies, (5) ribbon sandstone bodies, (6) stacked-bar sandstone bodies, (7) paleosol horizon, and (8) mudstone.
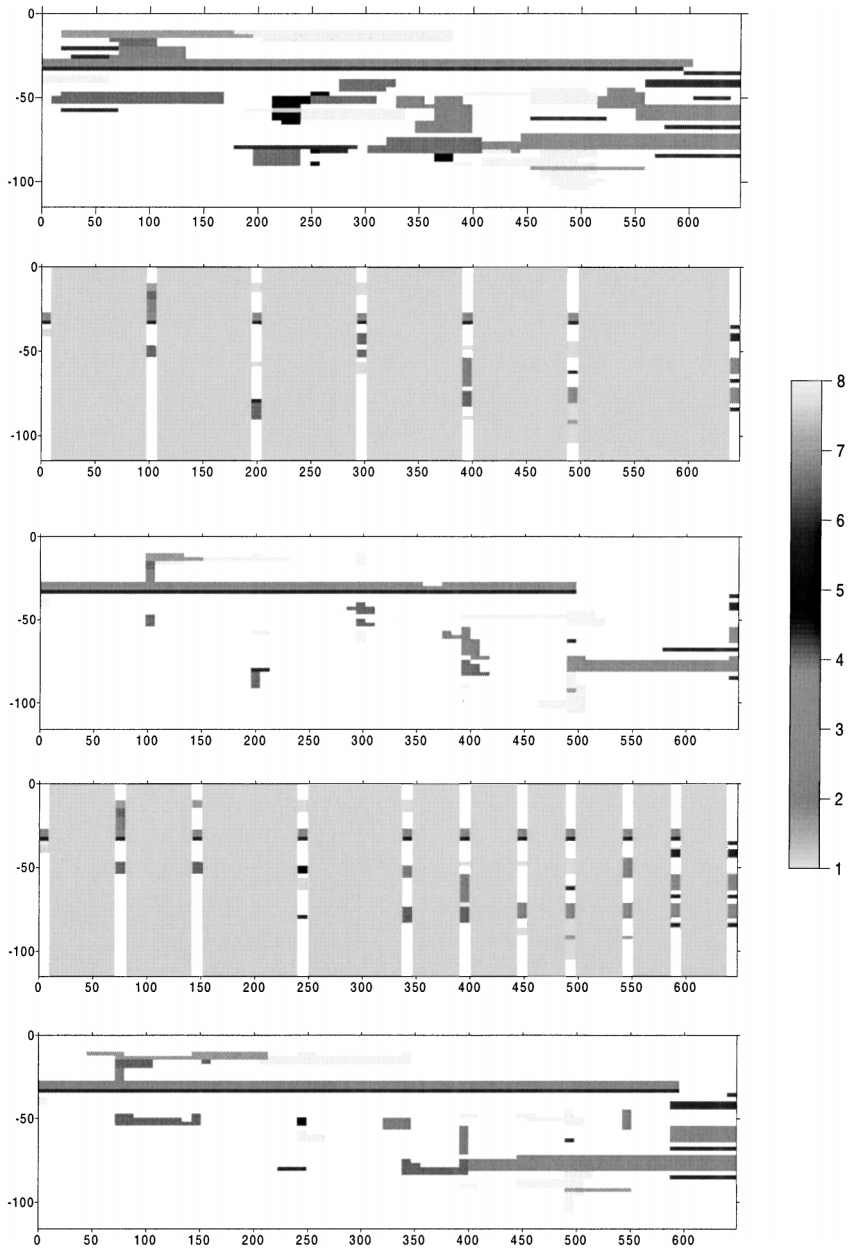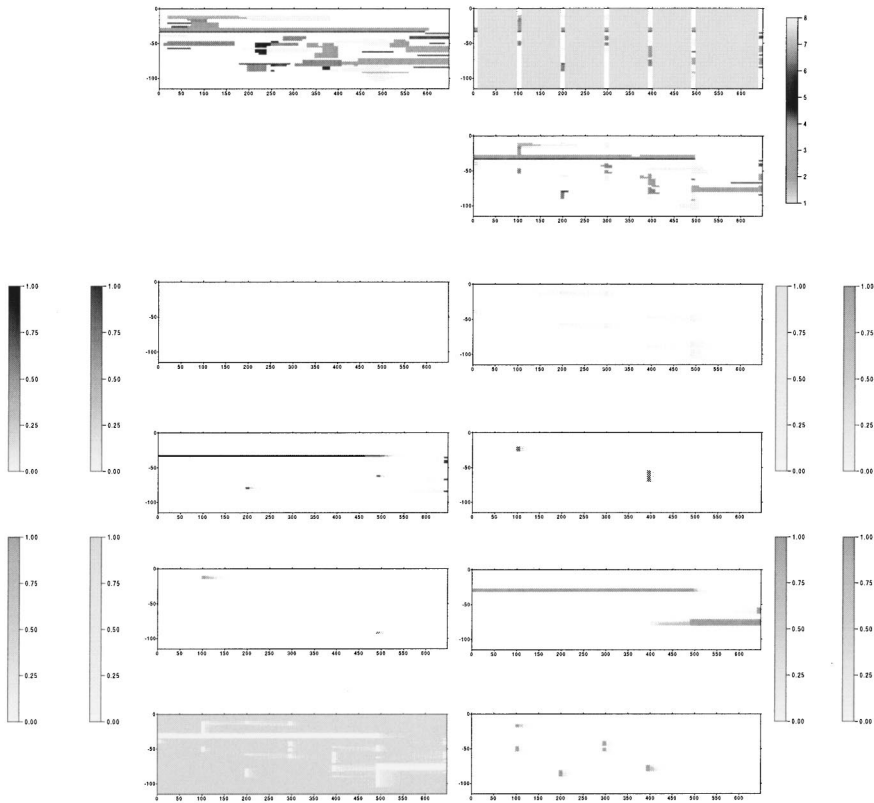
Table 3. Input Data for the Simulation of the Two-Dimensional Cross-Sectional Panel of the Fluvial Succession of the Medial Area of the Tórtola Fluvial System, Spain[a]

Length of the given section (m) = 648				Depth of the given section (m) = 115				
Sampling interval in X-axis (m) = 9				Sampling interval in Y-axis (m) = 2.5				
			No. of states = 8					

Horizontal transition probability matrix

State	1	2	3	4	5	6	7	8
1	0.893	0.009	0.005	0.000	0.000	0.000	0.000	0.093
2	0.000	0.796	0.011	0.000	0.000	0.000	0.000	0.194
3	0.000	0.000	0.989	0.000	0.000	0.000	0.000	0.011
4	0.006	0.000	0.013	0.885	0.000	0.000	0.000	0.096
5	0.074	0.000	0.000	0.074	0.593	0.037	0.000	0.222
6	0.000	0.013	0.000	0.000	0.000	0.946	0.000	0.040
7	0.040	0.000	0.000	0.000	0.000	0.000	0.940	0.020
8	0.007	0.006	0.002	0.007	0.005	0.005	0.001	0.968

Vertical transition probability matrix

State	1	2	3	4	5	6	7	8
1	0.591	0.000	0.000	0.000	0.014	0.000	0.042	0.353
2	0.011	0.753	0.097	0.000	0.000	0.000	0.000	0.140
3	0.032	0.000	0.623	0.000	0.000	0.238	0.000	0.107
4	0.000	0.025	0.000	0.662	0.013	0.000	0.000	0.299
5	0.111	0.000	0.000	0.074	0.519	0.000	0.000	0.296
6	0.000	0.000	0.026	0.032	0.006	0.084	0.000	0.851
7	0.120	0.000	0.000	0.100	0.000	0.000	0.360	0.420
8	0.029	0.008	0.039	0.017	0.003	0.031	0.010	0.863

[a]The states in the table are identified in Figure 9.

of the indicator function of each lithology is displayed in Figure 7. The same conclusions can be drawn as in the hypothetical case.

However, the geological features with short extensions are not very well reproduced. One of the advantages of this methodology is that, in conditioning to wells the geological features on a certain level (vertical coordinates) are kept at their level in the simulation. The object-based models used by Chessa and Martinius (1992) and Chessa (1995) do not have this advantage. Figure 6 (bottom row) shows the simulation results performed using the unconditional (on future states) coupled Markov chain model developed by Elfeki (1996). There are significant differences between the simulations that are performed with the conditional (on future states) and unconditional (on future states) coupled Markov chain models. Conditioning on future states is quite an achievement to make this methodology more practical.

Figure 9. Ensemble average over 100 realizations of the 2D cross-sectional panel of the fluvial succession of the medial area of the Tórtola fluvial system, Spain, shown in Figure 8. Top left is the schematic outcrop, top right shows the well locations; the second row is a single realization, from the third rows to the bottom are the ensemble averages of the indicators of each lithology. The grey scale ranges from 0 to 1.

Tórtola Fluvial System, Spain

Figure 8 (top image) shows the schematic picture of the two-dimensional cross-section panel of the fluvial succession of the medial area of the Tórtola fluvial system. The outcrop section shows the spatial distribution of the eight different genetic types that are distinguished and illustrated with different colors. The cross-section has a lateral extent of about 800 m and a stratigraphic thickness of about 115 m. The distance to the apex of the Tórtola fluvial system is

approximately 22 km. A detailed description of the Tórtola fluvial system can be found in Martinius (1996).

Stochastic simulation of the 2D cross-section of the outcrop was carried out. The statistical parameters that are used for the simulation are displayed in Table 3. These parameters are estimated from the schematic picture (Fig. 8, top image). Equation (24) was used to estimate the statistical parameters over a grid spacing of 9×2.5 m. Figure 8 shows the results of the stochastic simulation. The third and the fifth images are a single realization of the conditional simulation that is performed on the seven and eleven wells given in the second and the fourth images, respectively. The stochastic simulations in this example do not show significant differences between seven and eleven wells. This is due to the isolated geological features that are present in the outcrop. These features appear in one well and not in the others and are very sparse in the outcrop.

The ensemble average of the indicator function is also calculated and displayed in Figure 9. The sparse objects in this outcrop are also reflected in the ensemble average. It is also important to point out that the lithology coded 5 (black) does not appear in any of the wells and so it is reproduced neither in the single realization nor in the ensemble average.

CONCLUSIONS

An extension of the coupled Markov chain methodology developed by the first author has been performed. This methodology used information from a single well. It was not able to perform conditional simulations on more than one well. The extension, proposed in this paper, makes it more practical. Conditional simulations on any number of wells is now possible. The extension is based on the concept of conditioning Markov chains on the future states. A computer code called "SALMA" has been developed to implement the proposed methodology. The required input data for the program include the dimensions of the geological section (length and depth), the number of the geological materials present in the system, transition probabilities, sampling intervals over which these transitions are estimated, and well log data (the lithologies). The methodology has been tested on an artificially generated geological structure and on realistic outcrops. The methodology has proven fairly successful. The simulations presented in this paper use statistically homogeneous transition probability matrices. However, the methodology is flexible and can handle transition matrices that vary between the wells, or the case where the reservoir contains different large-scale layers and each layer has its own transition matrix. Generalization of the methodology will be considered in future work.

ACKNOWLEDGMENTS

comments and Prof. Ir. C. Van Kruijsdijk for giving the first author the opportunity to perform this research in the context of the current DIOC project (Multiscale Stochastic Modeling of Subsurface Heterogeneity). Thanks are also due to Dr. C. Kraaikamp for fruitful discussions and Dr. Kees Geel for providing the data of the outcrops.

REFERENCES

Billingsley, P., 1995, Probability and measure: third edition: Wiley—Interscience, New York, 515 p.

Carle, S. F., 1996, A transition probability-based approach to geostatistical characterization of hydros-tratigraphic architecture: unpubl. doctoral dissertation, University of California—Davis, 182 p.

Carle, S. F., and Fogg, G. E., 1996, Transition probability-based indicator geostatistics: Math. Geology, v. 28, no. 4, p. 453–477.

Chessa, A. G., 1995, Conditional simulation of spatial stochastic models for reservoir heterogeneity: unpubl. doctoral dissertation, Delft University of Technology, Delft University Press, Delft, The Netherlands, 174 p.

Chessa, A. G., and Martinius, A. W., 1992, Object-based modeling of the spatial distribution of fluvial sandstone deposits, *in* Christie, M. A., Da Silva, F. V., Farmer, C. L., Guillon, O., Heinemann, Z. E., Lemonnier, P., Regtien, J. M. M., and Van Spronsen, E., eds., Proc. 3rd Eur. Conf. on the Mathematics of Oil Recovery: Delft University Press, The Netherlands, p. 5–14.

Deutsch, C. V., and Journel, A. G., 1992, GSIB; Geostatistics Software Library and user's guide: Oxford University Press, New York, 340 p.

Elfeki, A. M. M., 1996, Stochastic characterization of geological heterogeneity and its impact on groundwater contaminant transport: unpubl. doctoral dissertation, Delft University of Technology, Balkema Publisher, The Netherlands, 301 p.

Galbraith, R. F., and Walley, D., 1976, On a two-dimensional binary process: Journal of Applied Probability, v. 13, p. 548–557.

Gozalo, M., and Martinus, A., 1993, Outcrop data-base for the geological characterization of fluvial reservoirs: An example from distal fluvial fan deposits in Loranca Basin, Spain, *in* North, C. P., and Prosser, D., eds., Characterization of fluvial and aeolian reservoirs: Geological Society Special Publication No. 73, London, p. 79–94.

Haldorsen, H. H., and Damsleth, E., 1990, Stochastic modeling: Journal of Petroleum Technology, v. 42, no. 4, April, p. 404–412.

Krumbein, W. C., 1967, Fortran computer programs for Markov chain experiments in geology: Computer Contribution 13, Kansas Geological Survey, Lawrence, KS.

Lin, C., and Harbaugh, J. W., 1984, Graphic display of two- and three-dimensional Markov computer models in geology: Van Nostrand Reinhold, New York, 180 p.

Martinius, A. W., 1996, The sedimentological characterization of Labyrinthine fluvial reservoir analogues: unpubl. doctoral dissertation, Delft, The Netherlands, 300 p.

Moss, B. P., 1990, Stochastic reservoir description: A methodology, *in* Hurst, A., Lovell, M. A., and Morton, A. C., eds., Geological applications of wireline logs: Geological Society of London, Special Publication 48, p. 57–75.

Neuman, S. P., 1980, A statistical characterization of aquifer heterogeneity: an overview, *in* Narasimhan, T. N., ed., Recent trends in hydrogeology, Special Paper Number 189, Geological Society of America, Boulder CO, p. 81–102.

Pickard, D. K., 1980, Unilateral Markov fields: Advances in Applied Probability, v. 12, p. 655–671.