Math Geosci (2008) 40: 133–146
DOI 10.1007/s11004-007-9142-4

Multiple-Point Simulations Constrained by Continuous Auxiliary Data

Tatiana L. Chugunova · Lin Y. Hu

Received: 8 August 2007 / Accepted: 3 December 2007 / Published online: 23 January 2008

Abstract An important issue of using the multiple-point (MP) statistical approach for reservoir modeling concerns the integration of auxiliary constraints derived, for instance, from seismic information. There exist two methods in the literature for these non-stationary MP simulations. One is based on an analytical approximation (the "τ-model") of the conditional probabilities that involve auxiliary data. The degree of approximation with this method depends on the parameter τ, whose inference is difficult in practice. The other method is based on the inference of these conditional probabilities directly from training images. This method classifies the auxiliary data into a few classes. This classification is in general arbitrary and therefore inconvenient in practice, especially in the case of continuous auxiliary constraints. In this paper, we propose an alternative method for performing non-stationary MP simulations. This method accounts for the data support in the modeling procedure and allows, in particular, continuous auxiliary data to be integrated into MP simulations. This method avoids the major limitations of the previous methods, namely the use of an approximate analytical model and the reduction of the auxiliary data into a limited number of classes. This method can be easily implemented in the existing MP simulation codes. Numerical tests show good performance of this method both in reproducing the geometrical features of the training image and in honouring the auxiliary data.

Keywords Geostatistical simulation · Multiple-point statistics · Training image · Data support · Non-stationarity

T.L. Chugunova · L.Y. Hu (✉)
Reservoir Engineering Division, IFP, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France
e-mail: l-ying.hu@ifp.fr

T.L. Chugunova
Centre de Géostatistique, Ecole des Mines de Paris, Paris, France

1 Introduction

Reservoir modeling based on multiple-point statistics is an active research topic in geostatistics. Unlike the traditional methods based on random functions, MP simulations use empirical multivariate distributions inferred from training images. This approach is flexible to data conditioning and allows reproducing certain complex architectures of geological facies. Initially, the MP approach was developed within a stationary framework (Guardiano and Strivastava 1993). However, in practice, the reservoir heterogeneity presents most often a trend in the reservoir field. This spatial trend is often accessible via information derived, for instance, from seismic attributes. Seismic information is of low (coarse) resolution but covers the whole reservoir field. Recent work on MP simulations accounts for auxiliary constraints (auxiliary data) derived from seismic (Strebelle 2000; Strebelle et al. 2002). In this paper, MP simulations of categorical variables (principal variable) constrained to a continuous auxiliary variable are considered. First, the existing methods for integrating auxiliary data into MP simulations are reviewed and then an alternative method to avoid some limitations of the existing ones is proposed.

2 A Review of the Existing Methods

An MP simulation is performed sequentially point by point in the simulation grid. Let u_0 be a simulation point with its closest neighborhood points $v_1, v_2, \ldots, v_n$. Consider the simulation of the principal variable Z at a point u_0, which takes K possible values or states. Suppose that some of the neighborhood points are already simulated and we have

$$Z(u_1) = z_1, \qquad Z(u_2) = z_2, \ldots, \qquad Z(u_m) = z_m; \quad \forall i, z_i \in (1, \ldots, K),$$

where $(u_1, u_2, \ldots, u_m) \subset (v_1, v_2, \ldots, v_n)$ and $m \leq n$. In the following, the complete neighborhood $v_1, v_2, \ldots, v_n$ is denoted by $\vec{v}^n$, the incomplete one $u_1, u_2, \ldots, u_m$ by $\vec{u}^m$, or simply $\vec{u}$, and the ensemble of simulated values $Z(u_1) = z_1, Z(u_2) = z_2, \ldots, Z(u_m) = z_m$ by $Z(\vec{u}) = z(\vec{u})$. In addition, at each simulation point u_0 we have auxiliary information (a constraint) derived, for instance, from seismic attributes. We denote this auxiliary variable as $S(u_0) = s, s \in \mathbb{R}$. The simulation of $Z(u_0)$ conditioned to $Z(\vec{u}) = z(\vec{u})$ and $S(u_0) = s$ requires the evaluation of the following conditional probability distribution

$$P\big[Z(u_0) = z | Z(\vec{u}) = z(\vec{u}), S(u_0) = s\big]. \tag{1}$$

For simplicity, this probability will also be written $P[Z(u_0) = z | z(\vec{u}), s]$, or $P(A|B, C)$, where A, B and C stand for events $\{Z(u_0) = z\}$, $\{Z(\vec{u}) = z(\vec{u})\}$ and $\{S(u_0) = s\}$, respectively.

Two methods for evaluating this probability distribution exist in the literature. The first method consists in using, in addition to one training image of the principal variable, a training image of the auxiliary variable (Strebelle 2000). The latter is obtained,

for instance, by numerical simulation on the first. We scan simultaneously both training images using the complete neighborhood $\vec{v}^n$ as a template, and we identify all patterns in the principal training image that satisfy (i) $Z(\vec{u}) = z(\vec{u})$, namely, the pattern of the principal training image at the already simulated neighborhood points is $z(\vec{u})$; (ii) $S(u_0) = s$, namely, the value of the corresponding auxiliary training image at the simulation point equals s. From this ensemble of patterns of the principal training image, we compute an empirical approximation of the conditional probability distribution (1). This direct and consistent method is, however, not practical, because of the often continuous character of the auxiliary variable. Indeed, there is generally very little chance to find a pattern, in both training images, which coincides at the same time with the given pattern of the given point neighborhood and with the given value of the auxiliary variable. A solution to this problem is to classify the auxiliary variable into a few classes. This may attenuate the above problem, but the reproduction of auxiliary characteristics is not continuous. In addition, the classification is in general somewhat arbitrary.

The second method is based on an analytical approximation of the probability (1) using the "τ-model" (Journel 2002). Based on this model, the probability of A jointly conditioned on B and C becomes a function of the marginal probability of A and the two probabilities of A separately conditioned on B and on C, i.e.,

$$P(A|B,C) = f_\tau\big[P(A),\, P(A|B),\, P(A|C)\big].$$

This function depends on a parameter τ that measures the degree of redundancy between events B and C with respect to event A (Krishnan et al. 2005). The quantification of this redundancy is very difficult in practice, which makes the inference of τ delicate. Moreover, this needs to be performed for each simulation point. For these reasons, the value of τ is set to 1 in the current use of this method. This corresponds to the case where B and C are conditionally independent given A. The degree of approximation of this choice is not easily appraisable and depends on each particular case. Nevertheless, when there are no relevant training images for both principal and auxiliary variables, the "τ-model" with an empirical τ value remains a practical solution for non-stationary MP simulations.

In this paper, we consider the case where both principal and auxiliary training images are available. For instance, a process-based simulation (Hu et al. 1994; Lopez 2003) can build the principal training image, and a subsequent process of geophysical modeling can build the corresponding auxiliary training image. Having both training images, we propose an alternative method to avoid the limitations of the above methods, namely the reduction of the auxiliary data into a limited number of classes for the first method, and the use of an approximate analytical model for the second method.

3 An Alternative Method

3.1 Principle

Consider that the auxiliary variable $S(u_0)$ reflects in general the average property of the principal variable $Z(\cdot)$ in a certain neighborhood of u_0, called the support of

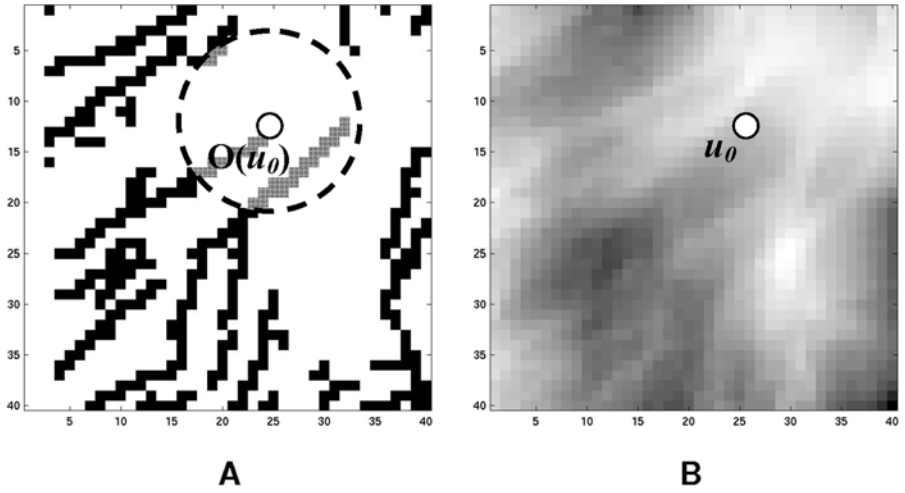

Fig. 1 Example of the support of an auxiliary variable: (**A**) Principal variable Z representing the spatial distribution of two facies; (**B**) Auxiliary variable S representing the proportion of the black facies on a disk

the auxiliary variable. S is often a smooth (or continuous) function in space. For instance, if the auxiliary variable represents the local orientation of a geological facies pattern, then its support is necessarily a domain instead of a single point. Another example is the probability of presence of a geological facies at a location u_0, that can be assimilated to the proportion of the facies in a certain volume (support) around u_0 (Fig. 1). In a grid system, such a support is discretized to a set of points, denoted by $\vec{v}^L = (v_1, v_2, \ldots, v_L)$, where L is the number of points reflected by the auxiliary property S (local proportion of facies, local size and orientation of patterns, etc.). Correspondently, $Z(\vec{v}^L)$ denotes a pattern of the principal variable on $\vec{v}^L$. For example, in the case of two facies ($Z = 0$ or 1), the local proportion S of facies 1 is simply $S(u_0) = \frac{1}{L}\sum_{i=1}^{L} Z(v_i)$.

Assume that we know the support of the auxiliary variable $\vec{v}^L = (v_1, v_2, \ldots, v_L)$. We use this support as a template to scan the two training images and to identify all patterns of the principal variable defined on the support $\vec{v}^L$. Each pattern contains a certain number (at least one) of replicates whose corresponding auxiliary data allow to build an empirical distribution of the auxiliary variable conditioned on this pattern $P[S < s|z(\vec{v}^L)]$. For simulating $Z(u_0)$ conditioned on $Z(\vec{u}) = z(\vec{u})$ and $S(u_0) = s$, we build an empirical approximation of the conditional distribution (1) in the following way. We retain all patterns defined on the support $\vec{v}^L$ that are compatible with $z(\vec{u})$. Then, among the retained patterns, we select those whose empirical distribution $P[S < s|z(\vec{v}^L)]$ is "compatible" with the auxiliary datum s. On the basis of these patterns, we build an empirical approximation of (1) that is subsequently used for drawing a value of the principal variable.

A criterion for the compatibility between the empirical distribution $P[S < s|z(\vec{v}^L)]$ and the auxiliary datum s will be proposed in the next section devoted to the implementation of the above alternative method. Note that in comparison with both existing methods, we introduced the notion of support of the auxiliary variable into

the alternative method, and the conditional probability of $Z(u_0)$ knowing $S(u_0) = s$ is not used. The example in Sect. 4.3 shows that the single auxiliary datum on the support $\vec{v}^L$ does not provide any significant information about the principal variable at the point u_0.

3.2 Practical Implementation

We choose to implement the above method into the currently used MP simulation approach based on a search tree structure for the storage of the MP statistics from training images (Strebelle 2000). As for the criterion of compatibility between an auxiliary datum and an empirical distribution, we simply check whether an auxiliary datum is between the minimal and the maximal values of an empirical distribution. Our future research will investigate other criteria of compatibility. The complete simulation procedure is composed of two main steps.

The first step concerns the scanning of the training images and the storage of the MP statistics. Consider that we have a training image of the principal variable and its corresponding training image of the auxiliary variable. We scan simultaneously both training images using $\vec{v}^L = (v_1, v_2, \ldots, v_L)$ as a template of size L, and we use the search tree structure to store the empirical MP statistics from the training images. This tree consists of an ensemble of nodes from level 0 to level L. The node at level 0 corresponds to the case where the neighborhood of the simulation point is empty. A node at level l (between 1 and L), denoted by $\vec{v}^l = (v_1, v_2, \ldots, v_l)$, is composed of l points closest to the simulation point. Each node of the tree corresponds to the particular pattern of the principal variable $z(\vec{v}^l)$. It is associated with an empirical probability distribution of the principal variable at the simulation point of this neighborhood given the pattern in this neighborhood $P[Z(u_0) = z|z(\vec{v}^l)]$ and an empirical probability distribution of the auxiliary variable given the same pattern

$$F_{z(\vec{v}^l)}(s) = P\big[S(u_0) < s|z(\vec{v}^l)\big]. \tag{2}$$

Like the empirical probability distribution of the principal variable, the empirical probability distribution of the auxiliary variable given $z(\vec{v}^l)$ is non-parametric and entirely defined by the auxiliary values corresponding to $z(\vec{v}^l)$.

The second step concerns the retrieval of conditional distributions and the simulations of the principal variable. The simulation is sequentially performed on a grid system according to a predefined path of visiting points. For simulating the principal variable at a point u_0, a value is drawn from the probability distribution (1) inferred in the following way:

(1) Determine the minimal level l of the tree nodes that include the already simulated neighborhood points $\vec{u}^m$, that is $l = \min_{k=1,\ldots,L}\{k : \vec{u}^m \subset \vec{v}^k\}$.
(2) Among all nodes at level l of the tree, identify those that contain the pattern $z(\vec{u}^m)$, that is $\{z(\vec{v}^l) : z(\vec{u}^m) \subset z(\vec{v}^l)\}$.
(3) Among these nodes, retain those whose auxiliary empirical distribution (2) is compatible with the auxiliary datum $s(u_0)$, that is

$$\big\{z\big(\vec{v}^l\big) : z\big(\vec{u}^m\big) \subset z\big(\vec{v}^l\big), s(u_0) \in \big[F^{-1}_{z(\vec{v}^l)}(0), F^{-1}_{z(\vec{v}^l)}(1)\big]\big\}.$$

4) Retrieve the MP statistics of such retained nodes and calculate the empirical conditional distribution (1).

4 Examples

The following simple examples illustrate the difference between the proposed method and the existing ones, and also show its applicability to diverse types of auxiliary constraints. For simplicity, there is no hard primary data conditioning in all the given examples.

4.1 Example of Two Facies Simulation

Consider first an example with two facies. The training image (Fig. 2A) is built by truncating a continuous Gaussian simulation. The truncation threshold is determined by the proportion function of the facies in black that varies linearly between 0 and 1 from the top to the bottom of the image (Fig. 2B). The simulation domain is composed of 100×200 grid blocks. The variogram of the Gaussian random function is of Gaussian type with a range of 3 grid blocks. We perform MP simulations using this training image and using the proportion function in Fig. 2B as an auxiliary constraint.

Using a template of 101 points, we scan simultaneously the training image on facies (principal variable) and the proportion function of the black facies that constitutes the training image of the auxiliary variable. We build a tree to store the multiple-point statistics obtained from the training images. Then we generate MP simulations using the above three methods for integrating the auxiliary data. Figures 2C, 2D and 2E show three realizations using respectively the τ-model with $\tau = 1$, the classification of the proportion in three classes and the alternative method. Figure 3 compares the experimental proportion functions of the black facies of the different realizations in Fig. 2. When using the alternative method, we obtain a slightly better reproduction of the geometrical features of the training image of Fig. 2A while respecting the imposed trend of facies proportion of Fig. 3A. Instead of the simple auxiliary constraint in Fig. 2B, consider now a more complex constraint to the proportion of the black facies in Fig. 4C. The principal and the auxiliary training images remain unchanged (Figs. 4A and 4B). Figures 4D, 4E and 4F show three realizations using respectively the τ-model with $\tau = 1$, the classification of the proportion in three classes and the alternative method. The geometrical features of these realizations look different from each other, and it is not easy to distinguish visually which one reproduces better the geometrical features of the training image. Nevertheless, the experimental proportions of these realizations (Fig. 5) show that the alternative method produces a better reproduction of the proportion constraint than the two other methods.

4.2 Example of Fracture Network

Now, consider an example from a real fractured reservoir. The principal training image (Fig. 6A) is a geological model of large-scale fractures in the reservoir field. The image in Fig. 6B represents the local fracture density, which is not constant in the reservoir field. We use this local fracture density as the auxiliary training image. Figures 6C, 6D and 6E show three MP simulations using respectively the τ-model with $\tau = 1$, the classification of the density in three classes and the alternative method. Figure 7 compares the experimental fracture density maps of the different realizations in Fig. 6. We observe a better reproduction of the geometrical features of the

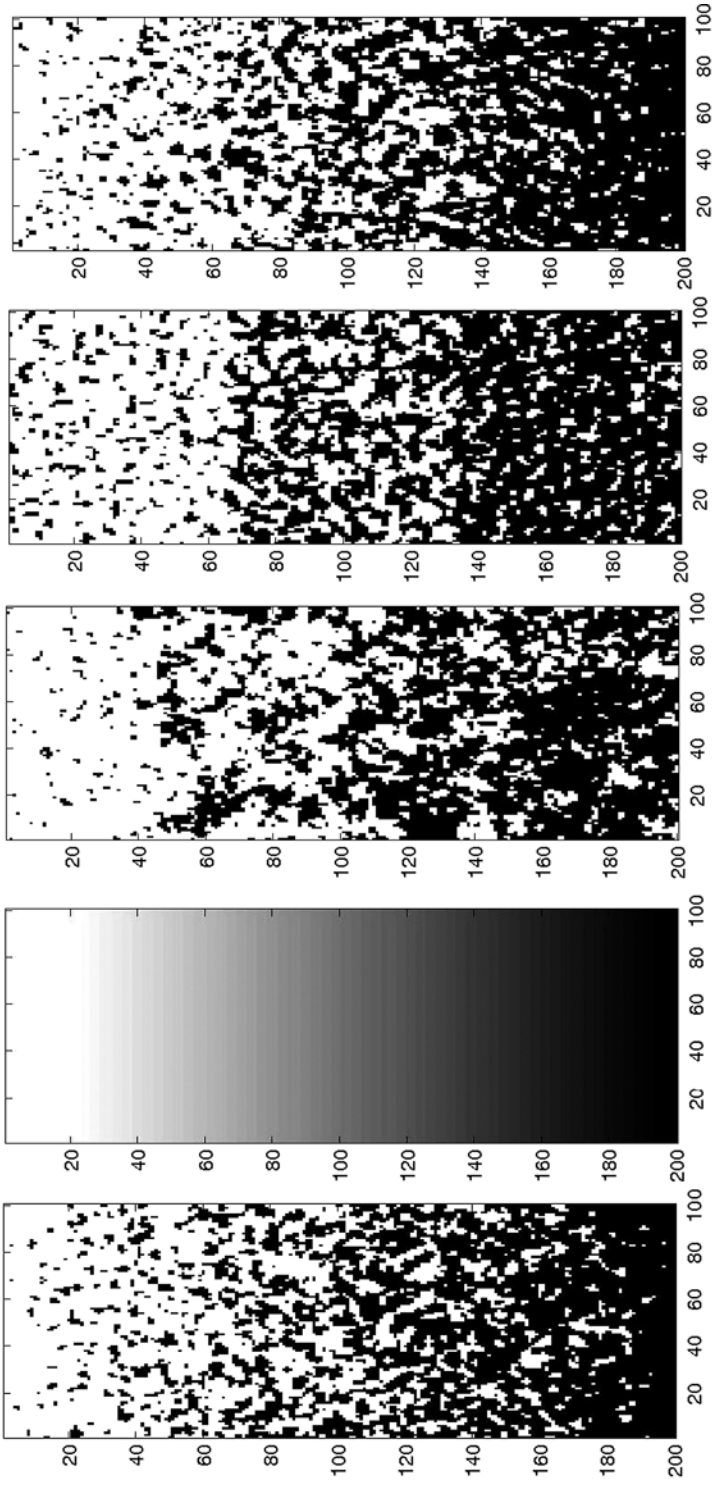

Fig. 2 Comparison of 3 methods of non-stationary MP simulation: (**A**) Principal training image; (**B**) Facies proportion used as auxiliary training image and as auxiliary constraint; (**C**) MP realization using the τ-model with $\tau = 1$; (**D**) MP realization using auxiliary data classification; (**E**) MP realization using the alternative method

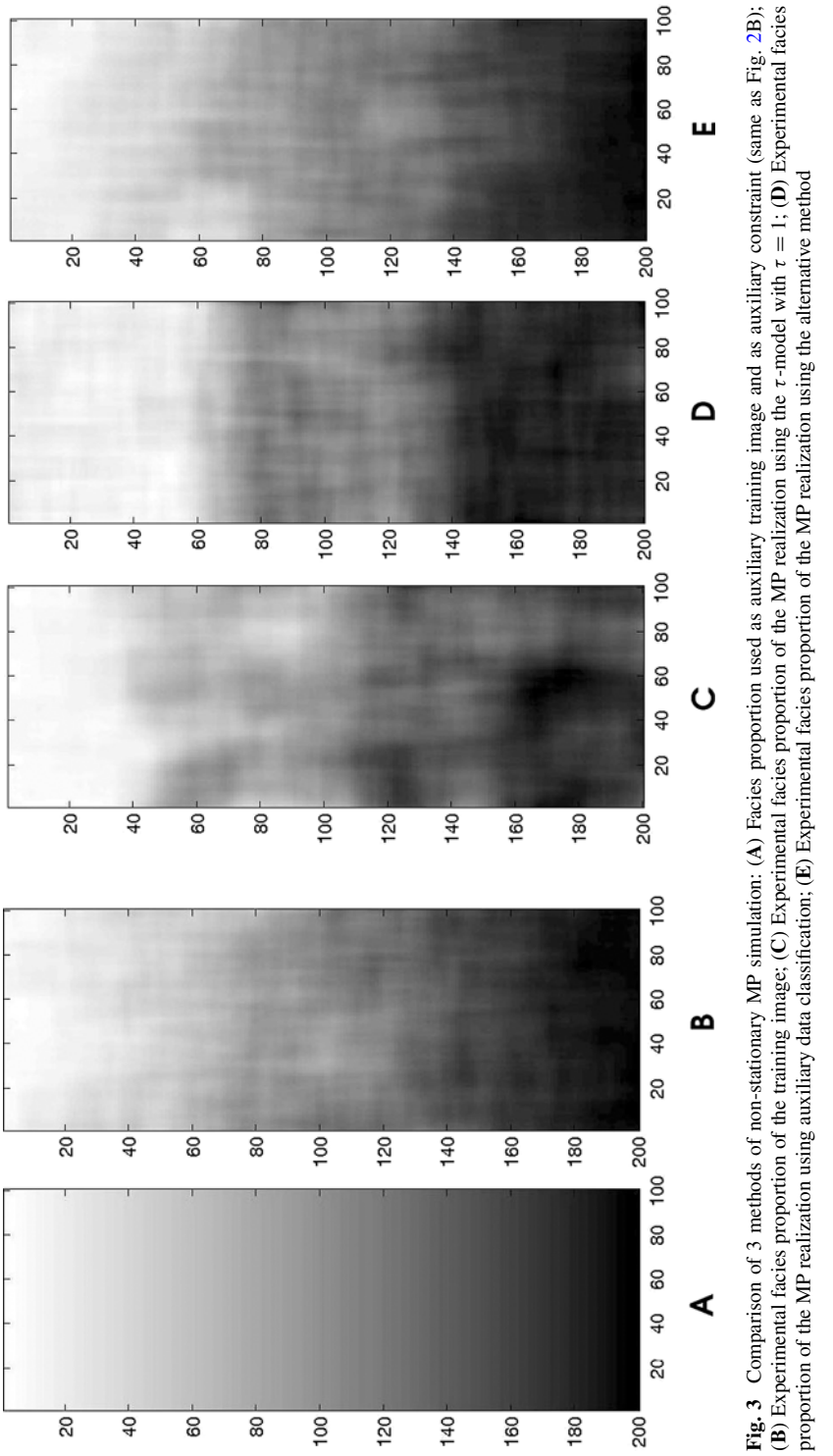

Fig. 3 Comparison of 3 methods of non-stationary MP simulation: (**A**) Facies proportion used as auxiliary training image and as auxiliary constraint (same as Fig. 2B); (**B**) Experimental facies proportion of the training image; (**C**) Experimental facies proportion of the MP realization using the τ-model with $\tau = 1$; (**D**) Experimental facies proportion of the MP realization using auxiliary data classification; (**E**) Experimental facies proportion of the MP realization using the alternative method

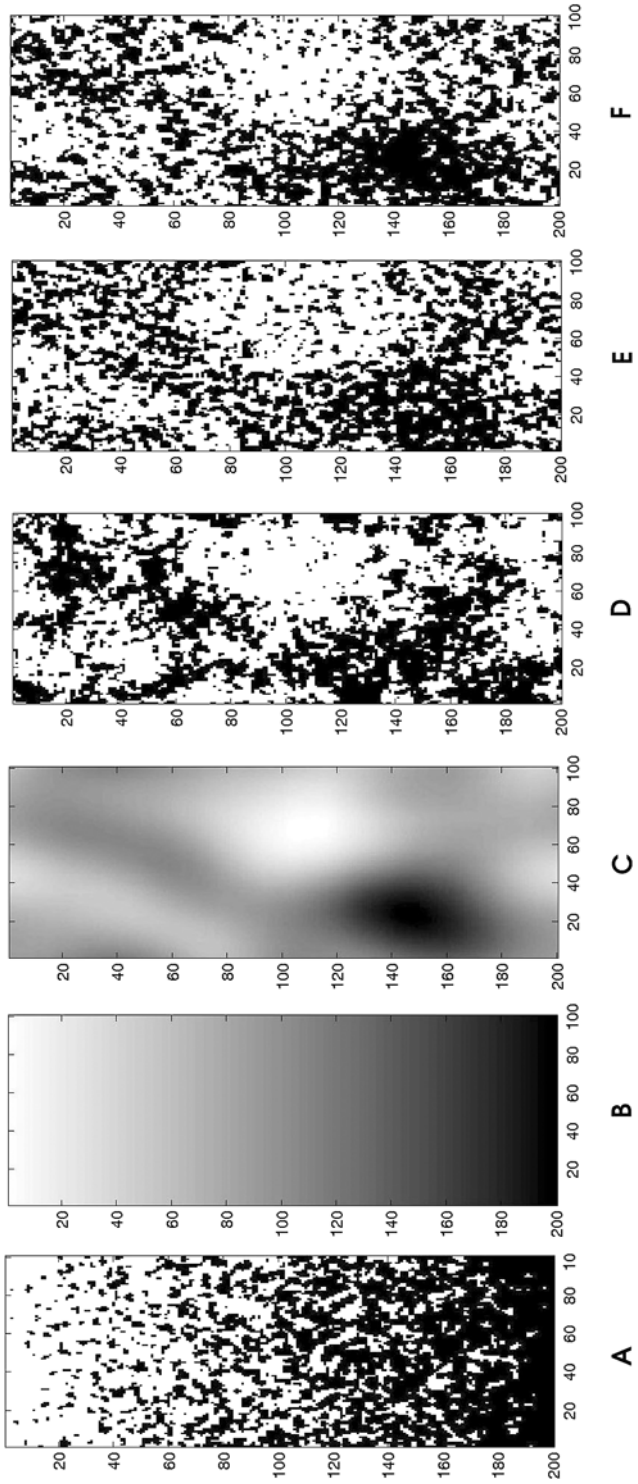

Fig. 4 Comparison of 3 methods of non-stationary MP simulation in the case of a more complex auxiliary constraint: (**A**) Principal training image (same as Fig. 2A); (**B**) Facies proportion used as auxiliary training image (same as Fig. 2B); (**C**) Facies proportion used as auxiliary constraint; (**D**) MP realization using the τ-model with $\tau = 1$; (**E**) MP realization using auxiliary data classification; (**F**) MP realization using the alternative method

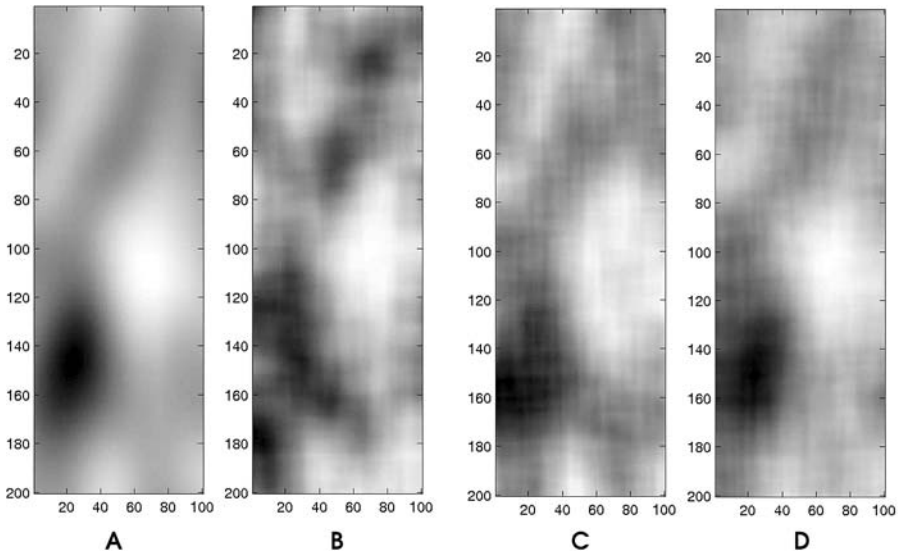

Fig. 5 Comparison of 3 methods of non-stationary MP simulation in the case of a more complex auxiliary constraint: (**A**) Facies proportion used as auxiliary constraint (same as Fig. 4C); (**B**) Experimental facies proportion of the MP realization using the τ-model with $\tau = 1$; (**C**) Experimental facies proportion of the MP realization using auxiliary data classification; (**D**) Experimental facies proportion of the MP realization using the alternative method

training image when using the classification of fracture density and the alternative method than when using the τ-model. Moreover, the experimental fracture density of training image (Fig. 7A) and that of the MP simulation using the alternative method (Fig. 7D) are very similar. The absolute difference of those density maps has the following statistics: mean $\sim 10^{-2}$ and variance $\sim 10^{-7}$.

4.3 Example of Orientation Constraint

In the above two examples, the auxiliary variable represents the local proportion of facies or fracture density. Consider an example where the auxiliary variable represents the orientation of local geometrical patterns. Figure 8A shows the principal training image obtained by the object-based simulation. The objects are randomly distributed in the simulation field with a constant density but orientated according to the auxiliary training image (Fig. 8B). Typically, the MP simulation using the τ-model (Fig. 8C) cannot reproduce this kind of auxiliary constraint. As a matter of fact, because the local orientation of objects (event C) does not provide any information about the local probability of object presence (event A), namely $P(A|C) = P(A)$, the τ-model, whatever the value of τ, leads to

$$P(A|B,C) = f_\tau\big[P(A), P(A|B), P(A|C)\big] = P(A|B),$$

where B stands for the principal conditioning data event.

However, because event C (object orientation) does affect event B (principal conditioning data event), namely $P(B|C) \neq P(B)$, we do not have $P(A|B,C) =$

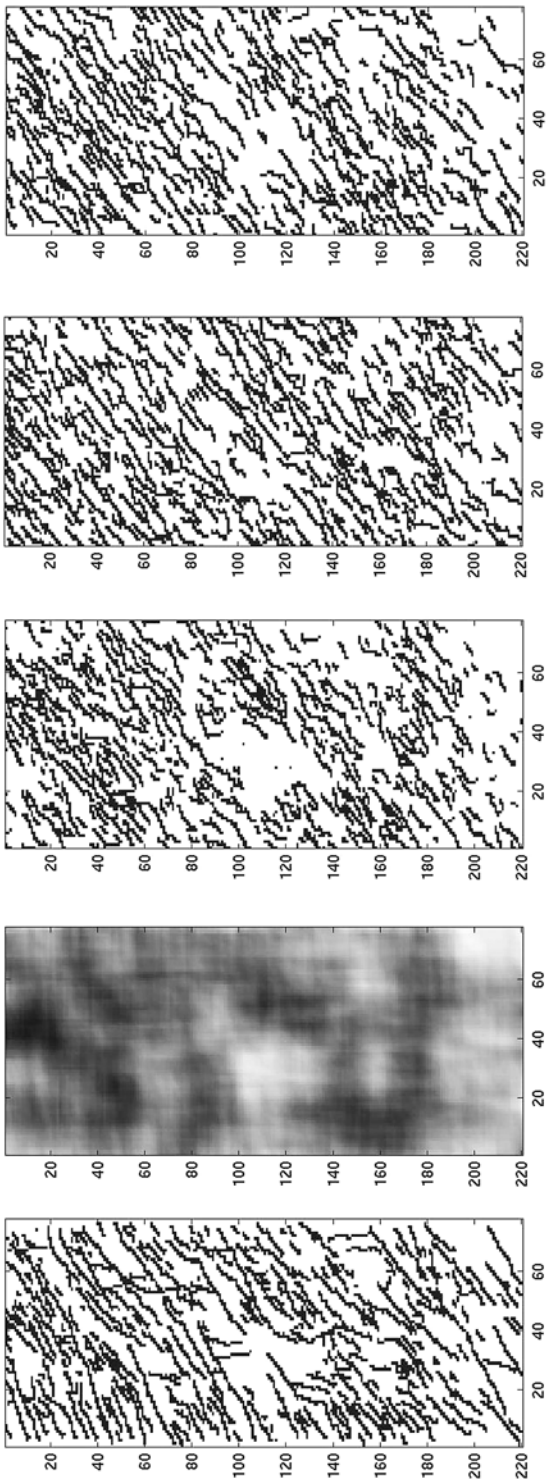

Fig. 6 Example of non-stationary MP simulation of fracture media constrained by local fracture density: (**A**) Principal training image; (**B**) Fracture density used as auxiliary training image and as auxiliary constraint; (**C**) MP realization using the τ-model with $\tau = 1$; (**D**) MP realization using auxiliary data classification; (**E**) MP realization using the alternative method

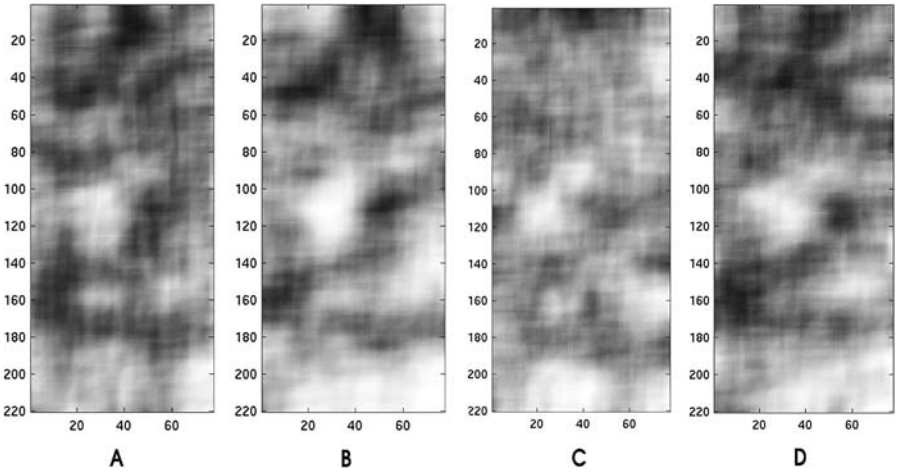

Fig. 7 Example of non-stationary MP simulation of fracture media constrained by local fracture density: (**A**) Fracture density used as auxiliary constraint (same as Fig. 6B); (**B**) Experimental fracture density of the MP realization using the τ-model with $\tau = 1$; (**C**) Experimental fracture density of the MP realization using auxiliary data classification; (**D**) Experimental fracture density of the MP realization using the alternative method

$P(A|B)$ in general, even if $P(A|C) = P(A)$. The MP simulation using the classification of orientation constraint (Fig. 8D) performs better than that of the τ-model, although there are artefacts in the simulation due to the classification of object orientation. Other similar methods by classifying pattern orientation into different regions (Strebelle and Zhang 2005; Wu et al. 2007) are not tested in this work, but should lead to similar results. Without any classification of the continuous orientation variable, the MP simulation using the alternative method (Fig. 8) performs very well both in terms of reproducing the geometrical features of the principal training image and in terms of honouring the orientation constraint.

5 Conclusions

In this paper, a new method for performing non-stationary MP simulation is proposed. This method is based on the direct inference of the conditional probabilities from the joint training images of the principal and the auxiliary variables. These training images can be built using numerical forward modeling based on geological and physical laws. The proposed method introduces the concept of data support into the modeling procedure and allows, in particular, continuous auxiliary data (facies proportion, fracture density, object size, orientation, etc.) to be integrated into MP simulations. The proposed method avoids the major limitations of the existing methods, namely the use of an approximate analytical model and the reduction of the auxiliary data into a limited number of classes or regions. It can be easily implemented in the existing MP simulation codes. Numerical tests show good performance of this method both in reproducing the geometrical features of the training image and in honouring the auxiliary data. The example with the object orientation as auxiliary constraint shows

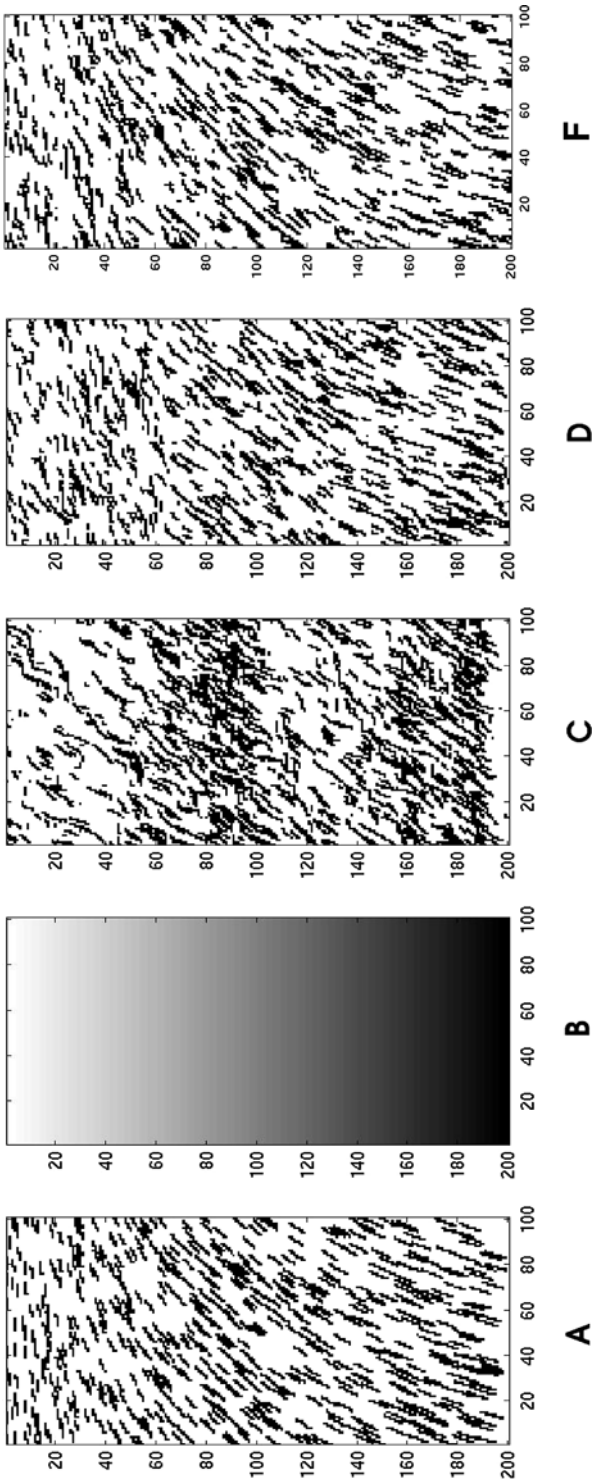

Fig. 8 Example of non-stationary MP simulation constrained by a map of object orientation: (**A**) Principal training image; (**B**) Orientation map used as auxiliary training image and as auxiliary constraint; (**C**) MP realization using the τ-model; (**D**) MP realization using auxiliary data classification; (**E**) MP realization using the alternative method

the particular interest of this method when the auxiliary variable does not provide significant information on the probability of presence of a facies at a single point, but reflects a geometrical characteristics of a facies pattern.

Acknowledgements The authors are grateful to Christian Lantuéjoul for valuable discussions and careful reading of the first draft of this paper. We also would like to acknowledge the two anonymous reviewers for their constructive and helpful comments.

References

Guardiano F, Strivastava M (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed) Geostatistics-Troia '92. Kluwer Academic, Dordrecht, pp 133–144

Hu LY, Joseph P, Dubrule O (1994) Random genetic simulation of the internal geometry of deltaic sand bodies. SPE Formation Eval, December 1994, 245–250

Journel AG (2002) Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses. Math Geol 34(5):573–596

Krishnan S, Boucher A, Journel AG (2005) Evaluating information redundancy through the TAU model. In: Leuangthong O, Deutsch C (eds) Geostatistics Banff 2004. Springer, Berlin, pp 1037–1046

Lopez S (2003) Modélisation de réservoirs chenalisés méandriformes: approche génétique et stochastique. PhD thesis, Centre de Géostatistique, Ecole des Mines de Paris

Strebelle S (2000) Sequential Simulation drawing structures from training images. PhD thesis, Petroleum Engineering Department, Stanford University

Strebelle S, Zhang T (2005) Non-stationary multiple-point geostatistical models. In: Leuangthong O, Deutsch C (eds) Geostatistics Banff 2004. Springer, Berlin, pp 235–244

Strebelle S, Payrazyan K, Caers J (2002) Modeling of a deepwater turbidity reservoir conditional to seismic data using multiple-point geostatistics. SPE 77425

Wu J, Zhang T, Boucher A (2007) Non-stationary multiple-point geostatistical simulations with region concept. In: Proceedings of the 20th SCRF meeting, Stanford, CA, USA