Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models

JEF CAERS AND TUANFENG ZHANG

Stanford University, Stanford Center for Reservoir Forecasting
Stanford, CA 94305-2220

January 2, 2002

Correspondence:

Jef Caers,
Stanford University
Petroleum Engineering Department
Stanford, CA, 94305-2220, USA
jef@pangea.stanford.edu
tel: 1-650 723 1774
fax: 1-650 725 2099

Abstract

While outcrop models can provide important information on reservoir architecture and heterogeneity, it is not entirely clear how such information can be used exhaustively in geostatistical reservoir modeling. Traditional, variogram-based geostatistics is inadequate in that regard since the variogram is too limiting in capturing geological heterogeneity from outcrops. A new field, termed multiple-point geostatistics does not rely on variogram models¿ Instead, it allows capturing structure from so-called ”training images”. Multiple-point geostatistics borrows multiple-point patterns from the training image, then anchors them to subsurface well-log, seismic and production data. Nevertheless, multiple-point geostatistics does not escape from the same principles as traditional variogram-based geostatistics: it is still a stochastic method, hence relies on the often forgotten principles of stationarity and ergodicity. These principles dictate that the training image used in multiple-point geostatistics cannot be chosen arbitrarily, and that not all outcrops might be suitable training image models. In this paper we outline the guiding principles of using analog models in multiple-point geostatistics and show that simple so-called modular training images can be used to build complex reservoir models using geostatistics algorithms.

Introduction

Over the last decade, geostatistics has been popularized as a quantitative tool for generating multiple reservoir models constrained to geologic, seismic and production data. Geostatistical reservoir characterization aims at achieving three goals

- Provide reservoir models that depict a certain believed or interpreted ”geological heterogeneity”;
- Provide a quantification of uncertainty through multiple reservoir models, all honoring that same geological heterogeneity;
- Integrate various types of data, each type bringing information on possibly different scales and with different precision.

Despite its widespread use, the application of geostatistics is hampered through the formal definition of ”geological heterogeneity”. As part of the mathematical sciences, geostatistical quantification and formalization of physical concepts is guided by strict mathematical rules and equations. This inevitably leads to a serious simplification or description of the natural processes or phenomenon under investigation. The champion of such simplification in geostatistics is the variogram, a measure of geological heterogeneity or continuity. The variogram is a statistical tool (not a geologic one) describing the dissimilarity of a variable observed at any two spatial locations. Successful application of the geostatistical methodology relies on the variogram, yet being a mere mathematical concept it has little connection to reality. Various strongly different types of reservoir heterogeneities may produce a similar experimental variogram, as shown in Figure 1.

The popularity of variogram-based geostatistics lies in the mathematical simplicity of the variogram model, not in its power to generate different types of geological models. Such mathematical convenience is developed in a ”geostatistics lab”, by geostatisticians who require a mathematical framework to make their estimation or simulation methods mathematically consistent. The concept of the variogram is, unfortunately, little motivated by actual geological considerations, and many practitioners have rightfully complained about its tedious interpretation and limited connection to actual geological phenomena. Outcrop data provide a rich source of geological structure in terms of faults, fractures, facies distribution and bedding configuration. The variogram is too limiting to capture this rich amount of information. Yet at the same time, a unified and flexible geostatistical methodology for capturing the geological richness of outcrop data is still lacking.

Before elaborating on our new geostatistical methodology, we will revisit the two basic, often forgotten, concepts in geostatistical reservoir modelling. First we discuss the principle of stationarity that requires ”similarity” and ”repetitivity” when pooling information together for the purpose of calculating univariate, bivariate (two-point) or multivariate (multiple-point) statistics. Next we elaborate on the concept of ergodicity which explain the inevitable fluctuation of these statistics and the consequence it has on model building. We will discuss these guiding principles from a non-mathematical point of view. Here we aim at presenting essentially the capabilities of current and future geostatistical methods and most importantly their limitations. Next, we present a novel approach to geostatistical reservoir modeling, termed multiple-point geostatistics. This methodology builds on the traditional two-point geostatistics (or variogram-based geostatistics), in the sense that its development strongly follows the two concepts outlined before. Multiple-point geostatistics has the potential to use outcrop data more completely than variogram-based geostatistics, to constrain better the reservoir heterogeneity. However, careful consideration is required on how to use outcrop data in geostatistics. This paper provides a guide to the successful use of such data in this new geostatistical approach to reservoir modeling. Rather than focussing on a particular case study, we present hypothetical examples that illustrate succinctly the points raised in this paper.

The geostatistical philosophy

Stationarity

Geostatistics relies on the well-known concept of random variables, in simple terms, reservoir properties at various grid location are largely unknown or uncertain, hence each property of interest at every grid block is turned into a random variable whose variability is described by a probability distribution function. In order to perform any type of geostatistical estimation or simulation one requires a *decision or assumption of stationarity*. *Any* statistical method, including geostatistics, relies on this assumption. Otherwise estimation of uncertain/random variables would not be possible. Various mathematical descriptions of this concept exist (Deutsch and Journel, 1998; Chiles and Delfiner, 1999). We will provide an intuitive one.

In order to determine the porosity histogram of a reservoir one relies on stationarity in the sense that information from various wells are pooled together into one single histogram. Implicitly, it is assumed that, in terms of a histogram, all the porosity values in that histogram originate from a single ”population”. The definition of population being a subjective choice itself. The population is often referred to as the zone of stationarity, a region that allows pooling information together. Hence any estimation of statistics such as mean and variance relies implicitly on a decision of stationarity. Such stationarity decision is not only relevant for simple statistics such as histograms, it carries over to higher order statistics. The variogram is a statistical measure of order two, since it describes the dissimilarity between the same (or different) variables at two spatial locations. Variograms are calculated by pooling information at similar lag distances together into a single scatter plot (a bivariate histogram essentially) from which the variogram value is calculated. Hence variogram calculations rely on a decision or assumption of stationarity.

It is important to list or make explicit the assumptions of stationarity. The model histogram and variogram are used for estimation and simulation purposes in all areas of the reservoir, often in areas where little a-priori information is present; We will see that a similar assumption of stationarity is required when exporting a variogram calculated from outcrop data to an actual reservoir.

While most users are well aware of stationarity decisions for the variogram and histogram, less attention is paid to the stationarity assumption of higher-order statistics of current geostatistical simulation methods. Indeed, one should accept that *any* mapping algorithm (also non-geostatistical) calls for the same amount of higher-order (or multiple-point) statistics. One cannot uniquely draw a map by knowing the histogram and variogram only, see Figure 1. Additional information in terms of higher-order statistics are needed. Take for example the well-known and popular sequential Gaussian simulation (sGs) algorithm. Most users are aware that only a variogram and histogram model are needed for this conditional simulation method to work. Nevertheless, hidden in the algorithm is the strong assumption that everywhere in the reservoir the higher-order/multiple-point statistics are multivariate-Gaussian. While such observation appears to be only relevant for a happy few mathematical geostatisticians, it has serious impact on the practice of reservoir modeling. The stationarity decision of multi-variate Gaussianity by definition generates reservoir models that are "homogeneously heterogeneous". The latter simply means that amongst all possible reservoir models that reproduce a given variogram model, sGs generates models that are maximally disconnected in the low and high values. This property of sGs is also termed the "maximum entropy" property. It is therefore troubling to observe that a large majority of reservoir models built today are based on assumptions that carry little geological relevance.

Geostatistics is driven, and at the same time limited, by a stationarity decision which is needed to make it work. Deciding on using geostatistics inherently includes a decision of stationarity. Essential parameters of the model are estimated based on a stationarity assumption. Algorithms are driven by stationarity assumptions since one relies on the fact the same algorithmic operation can be applied/repeated in every grid cell whose property requires estimation/simulation. We will develop a framework within the realm of multiple-point geostatistics where the strong stationarity assumption can somehow be relaxed.

Ergodicity

Geostatistical simulation algorithms reproduce input statistics such as a variogram and histogram under certain "ergodic fluctuations". These fluctuations are due to the limited, finite extent of the spatial domain being simulated. Sample statistics (net-to-gross ratio, porosity histograms, covariances) are not reproduced exactly in the sense that the variogram calculated from a single geostatistical realization does not match to the last digit, the input values of the variogram model. Such a match is not desirable anyway since sample statistics are inherently uncertain themselves. Ergodic fluctuations are not due to some geostatistical/mathematical trick that allows sample statistics to float due to their uncertainty. Simulation on an infinitely large domain will result in statistics of a realization that exactly match the model statistics. Therefore, when simulating on a finite domain, some statistics have smaller variations than other. For example the variogram statistics of a realization for small lag distances typically show less variability then the variogram for larger lag distances.

Ergodicity therefore plays an important role in both the estimation of model parameters as well as their simulation. It is typically advised in traditional geostatistical practice not to use any lag distance information beyond 1/2 the size of the field since they are not reliable (not enough samples to provide a reliable estimate of the variogram). Similar considerations will have to be made when using multiple-point geostatistics.

Multiple-point geostatistics

Variogram-free geostatistics

The application of variogram-based geostatistics in the petroleum industry is essentially carried over from its successful application in the mining world. Unfortunately, the purpose of utilizing geostatistics and the type of information available are quite different in mining applications than in oil/gas exploration and production. So too is the purpose of using geostatistics. While mining has abundant, so-called hard data (core measurement), oil and gas reservoirs are often typified by a lack of direct information. The purpose in mining is to estimate local recoverable grades of ore, while reservoir characterization aims at producing models that accurately predict global flow properties. Reservoir heterogeneity and accurate determination of the type of flow paths present are more important than a locally accurate quantification of permeability. Permeability models therefore do not have to be locally accurate, but do have to produce an accurate flow response, characterizing an integrated global measure of permeability heterogeneity.

The variogram cripples the petroleum geologist in her/his quest to produce accurate models of reservoir heterogeneity that may have serious impact on fluid flow. By describing merely correlations between only two spatial locations, a variogram cannot capture mathematically the complexity of curvilinear features (e.g. channels, cross bedding) nor can it describe any strong connectivities within a reservoir. Even if the variogram could be considered to be a reliable descriptor of reservoir heterogeneity, it is often very hard to determine its lateral component when only very few wells are available. The variogram provides an adequate description of geological heterogeneity within a single stratigraphic unit or facies. However, the definition of these stratigraphic units or facies models in three dimensions that has the largest impact on fluid flow in the reservoir.

One often refers to exhaustive data sets, such as outcrop analogs, to extract the required variogram parameters such as nugget, range and direction of anisotropy. But if such an abundance of information is available from outcrops, why restrict one selves to mere two-point correlations. Reproduction of curvilinear features from outcrops would require at least correlations between three spatial locations at a time and possible more, namely: multiple-point correlations or multiple-point statistics.

What is a training image ?

Multiple-point geostatistics relies on the concept of training images. Training images (the word image is slightly confusing because training images can be 3D) are essentially a database of geological *patterns*, from which multiple-point statistics, including the variogram, can be borrowed. Once the required patterns are extracted from the training image, they need to be anchored to subsurface data (e.g. well-log, seismic and production data). The variogram essentially is a statistical device to store patterns in a mathematical form, yet the complexity of the patterns is severely limiting: first, it is only two-point, secondly, its mathematical representation makes it hard to handle for non-experts. The training image replaces the variogram in multiple-point geostatistics as a measure for geological heterogeneity, it contains multiple-point information and, more importantly, is much more intuitive since one can observe, prior to any geostatistical estimation/simulation, what patterns will be reproduced in a set of multiple reservoir models.

Training images are often considered hard to come by; it appears in many cases easier to invent some variogram model than to define an entire complex 3D reservoir analog. Even if 3D exhaustive data is available, they are often considered too specific for the actual reservoir under investigation. Indeed limiting one selves to two-point statistics provides a false feeling of objectivity, namely, borrowing only two-points is less constraining than borrowing multiple-point statistics First of all, selecting a variogram model instead of a training image is not less ”subjective” or constraining. Recall the discussion above: any 2D or 3D modeling technique calls for the same amount of higher order statistics. If one chooses a variogram-based geostatistical method, the higher-order or multiple-point statistics are generated by the algorithm itself, hence hidden to the non-expert user, who might blissfully ignore the existence of any higher-order statistics. Training images allow an explicit quantification of geological heterogeneity prior to using any geostatistical method. It is easier to reject a training image based on visual inspection/interpretation for geological reality than to refute a variogram model or multi-Gaussian assumptions.

Nevertheless, similar to variogram estimation and modeling, the construction and use of training images are bound by the same principles of stationarity and ergodicity. Figure 2 shows three possible candidates for a training image: a set of elliptical shapes, a fluvial channel reservoir and a fan-like deposit. Under the principle of stationarity, only training image 2 can be used as a training image. Recall that training images are essentially databases to store patterns. If patterns are to be extracted from it, we will need to rely, as any statistical method does, on the stationarity of that pattern over the entire domain covered by that image. In other words, enough repetitivity is required to estimate or extract a set of multiple-point statistics from the training image. It is clear that training image 1 does not follow the principle of stationarity in the sense that patterns are changing over the training image. Consistency of patterns/texture is required. Training image 3 appears to be reasonably stationary in terms of its pattern but a closer inspection reveals that channels are thinner in the SE corner than in the NW, also the angle at which channel features occur varies from 0^o azimuth to 90^o, hence non-stationarity is present in the anisotropy of local patterns.

The size of the training image is an important factor. More importantly, the relative size of the training image with respect to the largest feature/pattern to be reproduced in the actual reservoir model needs to be carefully considered before any simulation takes place. This size consideration is required under the principle of ergodicity. When estimating parameters from a finite domain or when simulating reservoir models on a finite region, any statistics, including multiple-point statistics tend to fluctuate and that fluctuation becomes larger as the distances over which these statistics are calculated becomes larger. For example reproduction of large scale patterns (long correlations) such as channels (training image 2), would require a large training image, at least double the reservoir size in the direction of the channels continuity is recommended. A small training image would lead to large fluctuations of large range correlations, hence channels will tend to break up if the training image is taken the same size or smaller than the actual reservoir size. If smaller features are to be generated, such as in training image 1, a much smaller training image is required. In such case the training image can have a size much less than the actual reservoir.

It appears that under the restriction of stationarity and ergodicity, training images would be even harder to come by as may have been originally perceived (Caers et al., 2000; Strebelle, 2002). It should be clear that mere outcrop models, photographs of present day deposits or depositional systems, or simple drawings cannot necessarily be used as training images if they do not follow these underlying principles. One would expect that many outcrops essentially contain non-stationary features (as do actual reservoirs). The sizes of channel features or a net-to-gross ratio can vary considerably in the vertical as well as horizontal directions. Deltaic or turbidite deposits are inherently non-stationary in terms of geological patterns due to the changing depositional systems over the scale of a reservoir system (e.g. from tidal to subtidal for deltas). Does this doom the widespread use of multiple-point geostatistics ? On the contrary, we will develop simple techniques that relieves us essentially of the task of having to build a complicated 3D training image from outcrop data. In fact, we will show that strongly heterogeneous models can be built from simple so-called ”training image modules”, in the same sense as complex variograms can be built from a set of basic variogram models (e.g. spherical, Gaussian, nugget, etc ..). Before we proceed with building non-stationary models, we review the core geostatistical algorithm of multiple-point geostatistics.

The ”snesim” algorithm

Suppose a training image is available, such as training image 2 in Figure 2. The question considered in this section is how to design a geostatistical algorithm that can reproduce the patterns of this training image and at the same time honor any well and seismic data. A practical pixel-based algorithm to achieve this goal is the *snesim* algorithm (single normal equation simulation, see Strebelle, 2000, 2001). Snesim is a sequential simulation algorithm, much in the style of well-known methods such as sequential Gaussian simulation and sequential indicator simulation (Isaaks, 1990; Gomez and Journel, 1990). It relies on the idea of simulating each grid cell facies or petrophysical property sequentially along a random path, where the simulation of cells later in the process is constrained by cells earlier simulated, along with well and seismic data. Recall that a generic sequential simulation algorithm proceeds as follows:

- Construct a fine 3D grid with well-data assigned to closest grid cells
- Define a random path
- *Until* each non-datum cell on the random path is visited

step 1. Search for closest nearby well data and previously simulated cells

step 2. Construct a probability model for the property to be simulated based on the data found in step 1. and possibly any seismic data

step 3. Draw an outcome from the probability model in step 2 and assign that value to the current grid cell

- *end* simulation

Traditional to geostatistics is to use some form of kriging in step 2 to determine parameters of the probability model. The type of kriging determins the specific type (e.g. Gaussian, indicator, etc..) of sequential simulation. Kriging relies on a variogram model inferred from the data. In the snesim approach one shortcuts this step entirely by directly determining the probability model in step 2 from the training image, no kriging or variograms are involved. The details of this procedure are outlined in Strebelle (2000, 2002). Conditioning to soft (seismic) or production data is discussed in Strebelle (2002), Caers et al. (2001), Caers and Srinivasan (2001), Caers (2002).

The algorithm allows the definition of various training images at multiple scales. Instead of having to build a large complex training images, the method allows the use of a set of training images, each defining a variability at a different scale. For example, in the large scale training image one can define the large scale variability of channels, while smaller scale training images determine local channel features such as point bars, levees or crevasse splays, see Strebelle and Journel (2001).

Figure 3 shows a 3D training image generated using a Boolean approach (example borrowed from Strebelle, 2000). It should be stressed that this 3D model is not constrained by any specific well or seismic data, it is unconditional reflecting prior statistics on channel parameters. Figure 4 shows a single realization constrained to 20 vertical well data and a given vertical proportion curve (not shown). The resulting model reflects the channel patterns well, honors exactly all well data and reproduces the vertical proportion curve. To generate this 100.000 cell model, the method takes approximately 20 seconds on a 1GHz PC.

Figure 5 demonstrates the impact of ergodic fluctuation on using the snesim program. A 250x250 training image is provided, from which a single realization for each a 250×250, 100×100, 75×75 and 75×250 reservoir is simulated using snesim. It appears that disconnected channels occur in the 250×250 realization and not in any of the other models. The training image size must be at least 2 times larger than the largest feature reproduced in the reservoir model. Since NS channel features (=channel thickness) occur repeatedly in the NS direction, the reservoir size can be of dimension 250 in the NS direction as illustrated in Figure 5(e), without any major channel discontinuities in the simulated reservoir model.

Building non-stationary models: affinity and rotation invariance

Multiple-point geostatistics relies on the same principle of stationarity. The estimation or construction of a stochastic model relies on the availability of "repeated pattern information". Nevertheless, actual reservoirs may contain many non-stationary elements such as vertical and areal proportion variations, or a change in shapes of important flow bodies. This section addresses the question of how to borrow patterns from a stationary training image, then construct non-stationary reservoir models from it. First we introduce the notion of rotation and affinity invariance.

Rotation and affine transformations

Traditional variogram-based geostatistics relies on the estimation and modeling of nested variogram models from basic variogram structures. Basic models such as a unit range Gaussian, spherical, exponential and hole effect model are appropriately rotated and affinely (linearly) transformed, then combined with an appropriate nugget effect to form a model to fit experimental data. A typical fitting procedure consists of first determining major anisotropy directions (azimuth, dip), then determining the magnitude of continuity (or range) along that direction for each basic component. Once these angle and affinity parameters are determined a suitable transformation is setup to change the Cartesian coordinate system accordingly. In its more general form a nested variogram model is written as:

$$\gamma(|\mathbf{h}|) = \sum_{k=1}^{K} \gamma(|\mathbf{h}_k|)$$

where each component has its own anisotropy structure $\mathbf{h}_k$ defined by a linear transformation:

$$\mathbf{h}_k = A\, R_\theta\, \mathbf{h}$$

The matrix R_θ contains the rotation angle azimuth θ (if we work in 2D for demonstration purposes):

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

The matrix A contains the major and minor range of continuity, a_x and a_y:

$$A = \begin{pmatrix} a_x & 0 \\ 0 & a_y \end{pmatrix}$$

Based on the same principles of rotation and affinity transformation, we can build a similar rich set of multiple-point structures. Nesting of models is obtained by using different training images for different scales of observations as outlined above. Using rotation and affinity transforms, we can start from a simple training image of rectilinear channels, then generate a large variety of channel-type reservoirs from it. For example, suppose a training image X of channels running East-West is available with channel thickness equaling 100m on average. The principle of rotation and affinity transformation allows generating channel reservoir Y using training image X, where Y has North-South channels of only 50m thickness. This is achieved through the following proposed modification of the "snesim" algorithm, see also Figure 6.

- Construct a fine 3D grid of well-data to closest grid cells
- Define a random path
- *Until* each non-datum cell $\mathbf{u}$ on the random path is visited

step 1a. Search for closest nearby well data and previously simulated cells. Call this set a "data event".

step 1b. Rotate and affinely transform the data event accordingly. The center of such transformation is the visited cell location $\mathbf{u}$ to be simulated.

step 2. From the training image, read a probability for the property to be simulated based on the rotated data from step 1b. and possibly any seismic data

step 3. Draw an outcome from the probability model in step 2 and assign that value to the current grid cell

- *end* simulation

The local data event is rotated in the opposite direction of the desired continuity. Then the facies probability model is determined from the original non-rotated training image. Note that this procedure does not rotate the training image itself, rather it rotates local data events along the path of sequential simulation. In mathematical terms, each single datum with original coordinates $\mathbf{u}^{orig} = (x^{orig}, y^{orig}, z^{orig})$ in the entire data event is rotated along the center node to new coordinates $\mathbf{u}^{new} = (x^{new}, y^{new}, z^{new})$ according to

$$\mathbf{u}^{new} = A\, R_\theta\, \mathbf{u}^{orig}$$

Figure 7 demonstrates the rotation and affinity transforms. Figure 7 provides the training image for each of the resulting reservoir models in Figure 7(b) to (f). Note that the training image is not rotated or affinely transformed itself, the EW trending channel patterns of the training image are used to construct various other channel features that have differing channel sinuousities, thicknesses and orientation. The training image can therefore be termed a *training image module*, a template image or basic image (like a basic variogram) whose patterns can be rotated and affinely transformed.

Local rotation and affinity variations

An even richer set of reservoir models can be generated from a single training image module, by allowing the rotation angle and affinity correction factor to vary locally. The snesim algorithm remains exactly the same, only one allows for local variability in

$$\mathbf{u}^{new} = A(\mathbf{u})\, R_{\theta(\mathbf{u})}\, \mathbf{u}^{orig}$$

where $\mathbf{u}$ are the coordinates of the center cell and

$$A(\mathbf{u}) = \begin{pmatrix} a_x(\mathbf{u}) & 0 \\ 0 & a_y(\mathbf{u}) \end{pmatrix} \quad R_\theta(\mathbf{u}) = \begin{pmatrix} \cos(\theta(\mathbf{u})) & -sin(\theta(\mathbf{u})) \\ sin(\theta(\mathbf{u})) & cos(\theta(\mathbf{u})) \end{pmatrix}$$

This procedure requires the local angle and affinity factors to be known in every grid cell. Very often such parameters can be easily determined or interpreted from seismic, well-data or geological information, or could be stochastic (unknown) by itself, hence their (two-point) statistics borrowed from outcrop data (Xu, 1996).

Figure 8 provides an example of this approach. Two maps of local rotation and affinity transform factors are provided in Figure 8(b) and (c), a training image of EW channels shown in Figure 8(a). Three realizations constrained to the local angle and affinity factors are shown in Figure 8(d-f).

It would be a mistake to use a training image of a distributed system, such as Figure 8(d) as a training image in the snesim algorithm. Figure 9 shows that the use of a ”non-stationary” training image leads to reservoir models that do not capture the training image structure.

Net-to-gross ratio variations

The snesim algorithm does not require the training image to have the same facies proportion as the actual reservoir, although reservoir and training image net-to-gross ratio should not be greatly different (e.g. 30 % versus 75 %). A so-called servo-system (see Strebelle, 2000, 2001) is in place which progressively corrects any difference in facies proportions between the target proportions and the training image facies proportions. In fact, the technique allows one to constrain reservoir models easily to target vertical proportion curves (see Strebelle and Journel, 2001), or areal proportion maps, obtained from seismic and production data (Caers and Srinivasan, 2001).

Comparison with Boolean models

Multiple-point geostatistics is an alternative to building reservoir models using object-based techniques (Deutsch & Wang, 1996; Holden et al., 1998). While object-based techniques are a considerable improvement over traditional variogram-based methods, they have some important shortcoming compared to multiple-point geostatistics, namely:

- Conditioning to dense data set is virtually impossible
- Conditioning to seismic data is usually limited to constraining to vertical proportion curves and areal proportion maps. Actual integration of the physics itself is not viable for large models within a reasonable CPU. Multiple-point geostatistics is much more flexible in integrating multiple seismic attributes, as well as accounting for the rock physics considerations (Caers et al. 2001).
- Runtimes can be large for large models with dense data sets.
- Require geological models to be in the form of objects. This might work well for channel-type reservoirs but is likely to work to a lesser degree for carbonates, where facies association are important and important architectural elements might not be easily parameterized by simple objects. Multiple-point geostatistics can integrate any type of geology as long as it can be represented as a training image or a combination of training image modules.

A final perspective

Probably the best way to construct a single geologically realistic reservoir model is to build it manually using ones' own expert knowledge. The problem with this approach is that it will provide only a *single* model and that integrating other types of data, such as seismic and production data, requires other expertise. Geostatistics essentially provides a reasonably automated approach to building multiple reservoir models integrating (compromising) between multiple types of information. This automation comes at the price of model simplification. Automated or stochastic methods rely on a model or algorithm that can produce multiple models without much user intervention. This requires the design of a model or algorithm that can capture the complexity of all available information. Such model necessarily relies on the principle of repetitivity or stationarity. The improvement of multiple-point geostatistics over traditional geostatistics lies in the capability of capturing complex heterogeneity. Nevertheless, the geostatistical guiding principles remain the same: if information is not repetitive (stationary) it cannot be captured and reproduced by a model that relies on stationarity.

Therefore not all outcrop models can be directly used in multiple-point geostatistics as a reservoir analog or training image, see Figure 9. Some outcrops might be too specific, i.e. do not contain enough stationary information/geological patterns to be captured by the snesim technique. This paper proposes the first steps to modular training images as to resolving this problem. Such training images contain regularly occurring patterns within outcrops (channel, ellipse, cross-bedding, fracture shapes). A combination (nesting) of, rotation and affinity transformations of training image modules facilitates the generation of a wide variety of complex structures. The remaining research challenge now lies in finding a systematic procedure for

1. Finding which are the most important modular training images. Analogous to the existing basic variogram models, a limited set of modular training images could allow capturing geologic variability within most existing reservoir types.

2. Combining/rotating/affine transforming modular training images (similar to variogram fitting). This could ultimately lead to a catalog of training images from which most reservoir could be modeled. to obtain a desired geologic model. In this paper we only discuss the possibility of using the basic and simple modular training images for generating complex reservoir reservoirs. We did not discuss or propose a technique for determining the parameters that define such transformations. In essence this problem is similar to determining the rotation and affinity parameters for a given experimental variogram model.

Although this methodology has been applied to actual reservoirs, we avoided any demonstration in order not to confuse the methodological and conceptual emphasis of this paper. in this regard, we refer to other recent (see references) and upcoming publications.

References

Caers, J., Srinivasan, S. and Journel, A.G., 2000. Geostatistical quantification of geological information for a fluvial-type North Sea reservoir. SPE Reservoir Eval. & Eng., v. 3, p. 457–467.

Caers J., and Srinivasan, S., 2001. Statistical pattern recognition and geostatistical data integration. In: Soft computing for reservoir characterization and modeling, eds. P. Wong et al., Springer Verlag, p. 355-386.

Caers, J., Avseth, P. and Mukerji, T., 2001. Geostatistical integration of rock physics, seismic amplitudes and geological models in North-Sea turbidite systems. The Leading Edge, v. 20, no. 3, p. 308–312, March 2001.

Caers, J., 2002. History matching under training-image based geological model constraints, submitted for publication.

Deutsch and Journel, 1998. GSLIB - Geostatistical Software Library and User's Guide, 2^{nd} edn. Oxford University Press, New York, 369pp.

Deustch, C.V. and Wang, L., 1996. Hierarchical object based stochastic modeling of fluvial reservoirs. Mathematical Geology, v. 28, p. 857–880.

Chiles, J.P. and Delfiner, P., 1999. Geostatistics: modeling spatial uncertainty. Wiley & Sons, New York.

Gomez, J. and Srivastava, R., 1990. ISIM3D: an ANSI-C three dimensional multiple indicator conditional simulation. Computer & Geosciences, 16, 395–410.

Holden, L., Hauge, R., Skare, O., Skorstad, A., 1998. Modeling of fluvial reservoirs with object models, Mathematical Geology, 30, 473-496.

Isaaks, E., 1990. The application of Monte Carlo methods to the analysis of spatially correlated data. PhD Thesis, Stanford University, Stanford, CA, USA

Strebelle S., 2000. Sequential simulation drawing structure from training images. PhD Thesis, Stanford University, Stanford, CA, USA

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology, v. 34, p. 1-22.

Strebelle, S. and Journel, A.G., 2001. Reservoir modeling using multiple-point statistics. SPE Annual Technical conference and Exhibition, New Orleans, Sept. 30 – Oct. 3, 2001, SPE # 71324.

Xu, W., 1996. Conditional curvi-linear stochastic simulation using pixel-based algorithms. Mathematical Geology, v. 28, p. 937–950.

3 different "heterogeneities"

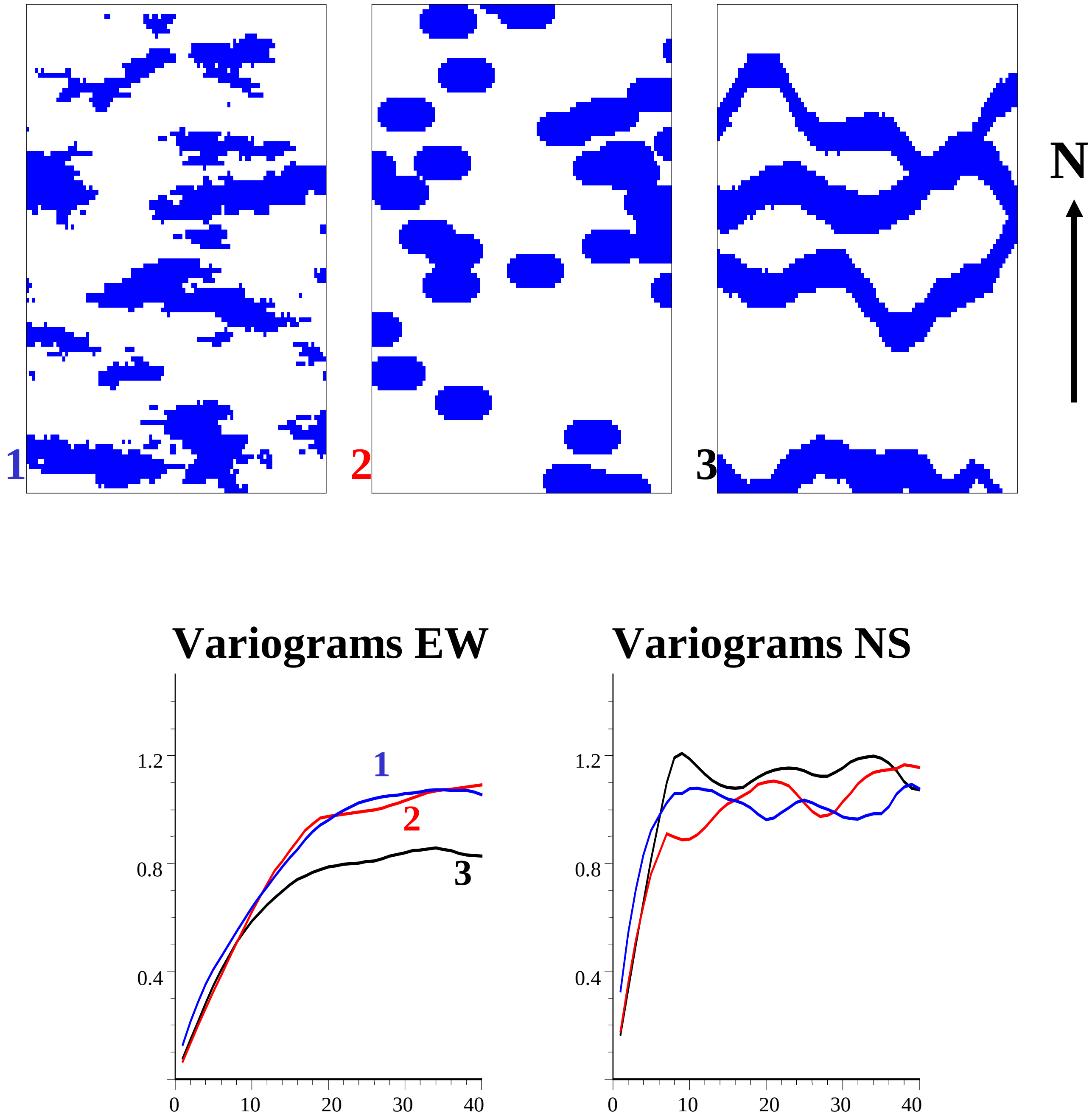

Figure 1: The variogram as a poor descriptor of geological heterogeneity. Three different geological heterogeneities result in three similar variograms

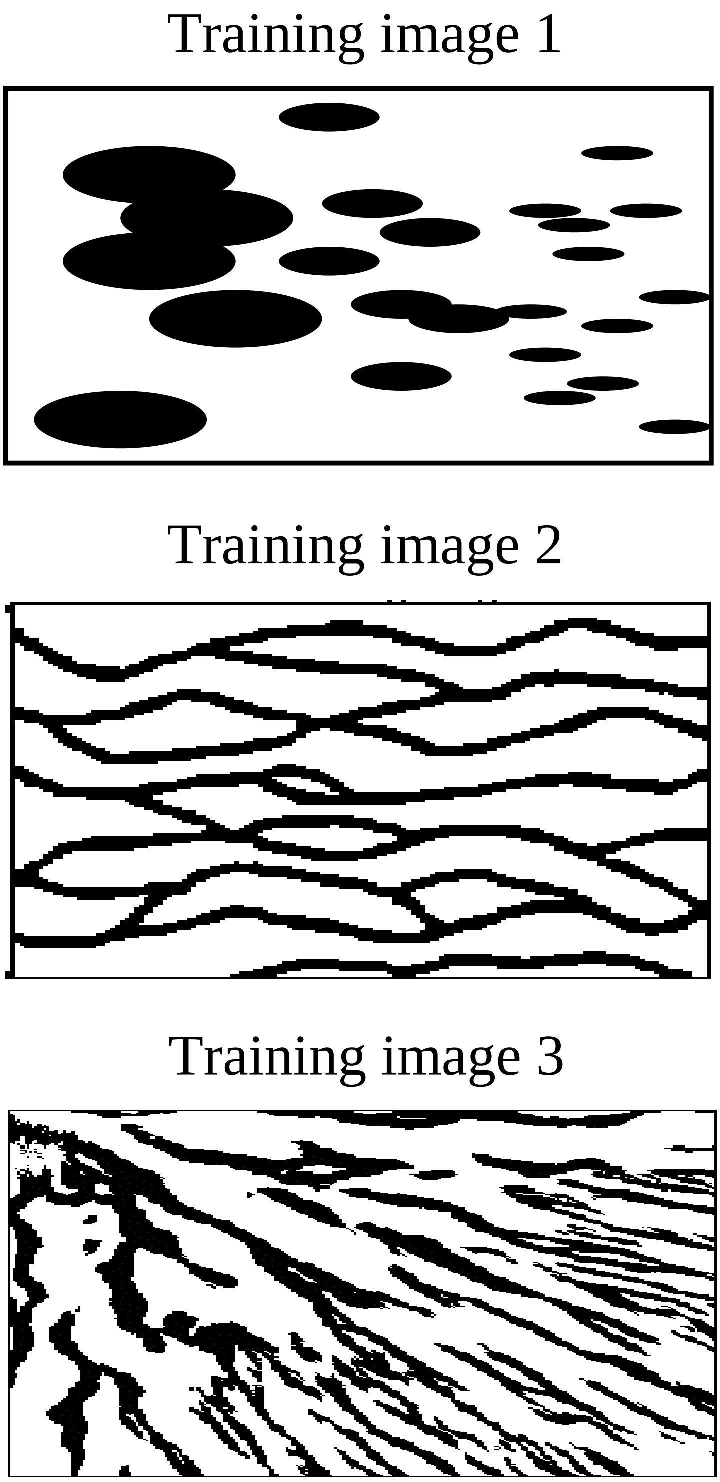

Figure 2: Three possible candidates for a training image. (a) elliptical shapes, (b) fluvial type reservoir, (c) deltaic type reservoir. Only image two can be used as a training image since pattern are stationary over the entire image.

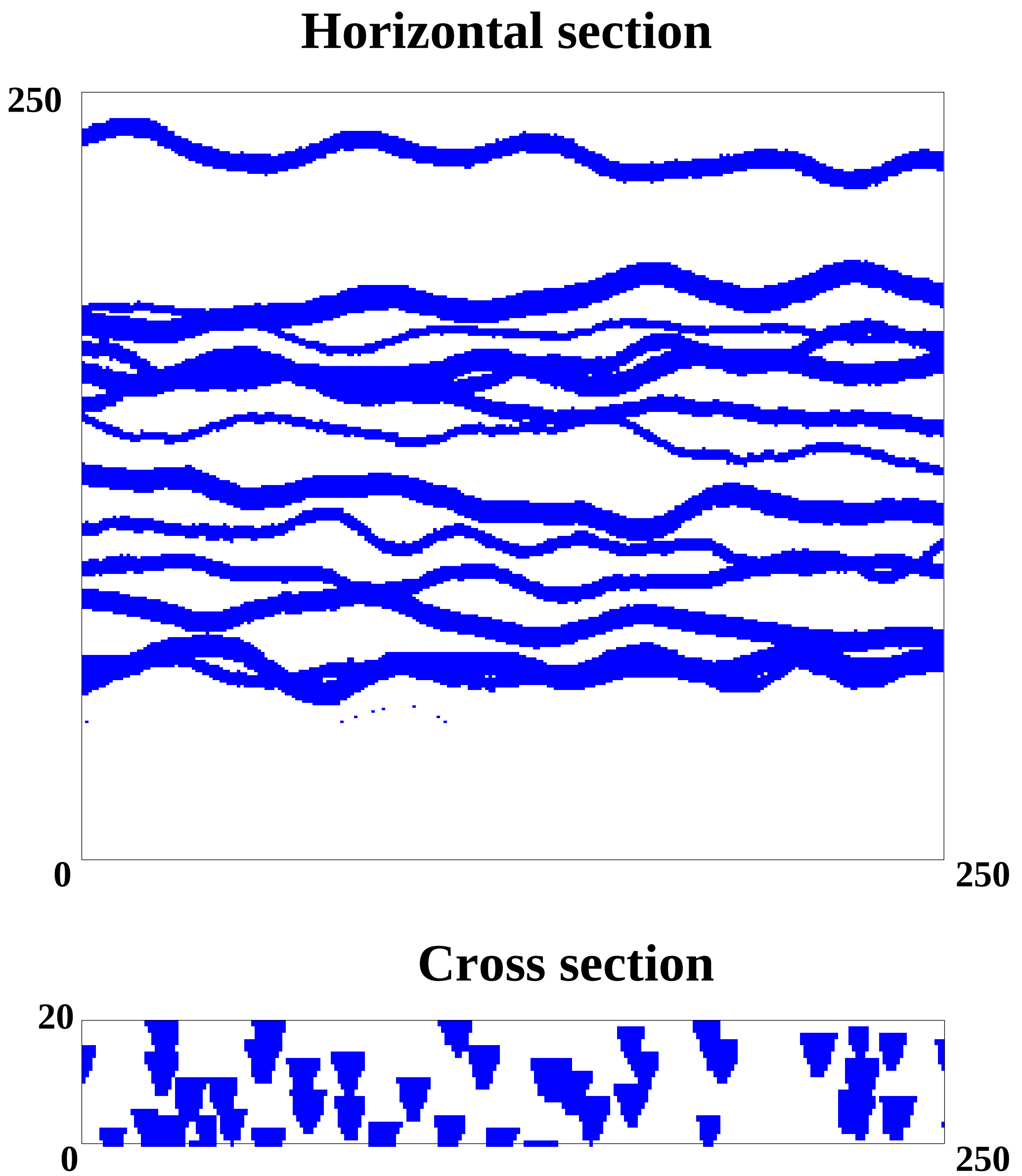

Figure 3: 3D training image generated from a Boolean method

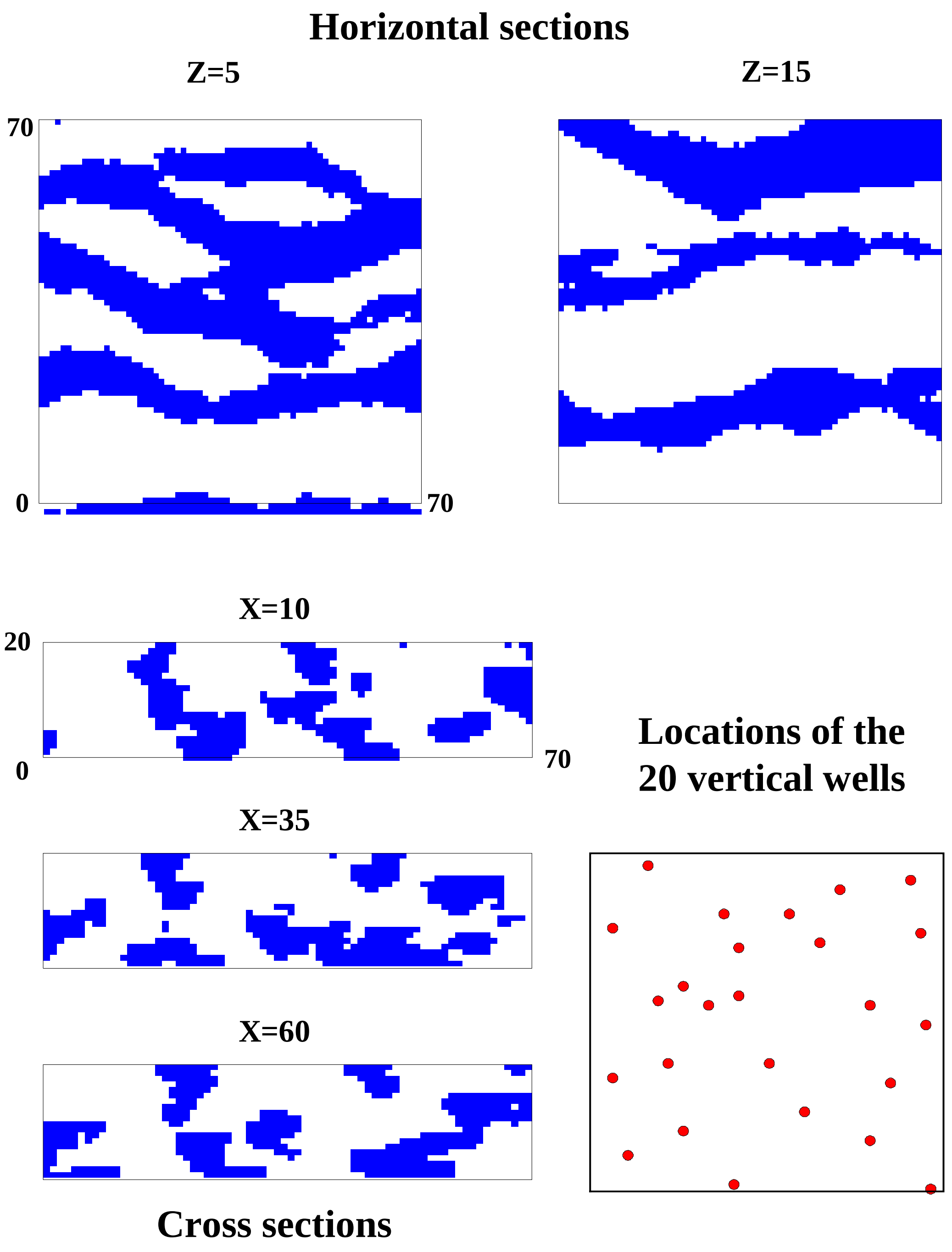

Figure 4: Single snesim realization constrained to 20 vertical well data, using the training image of 3

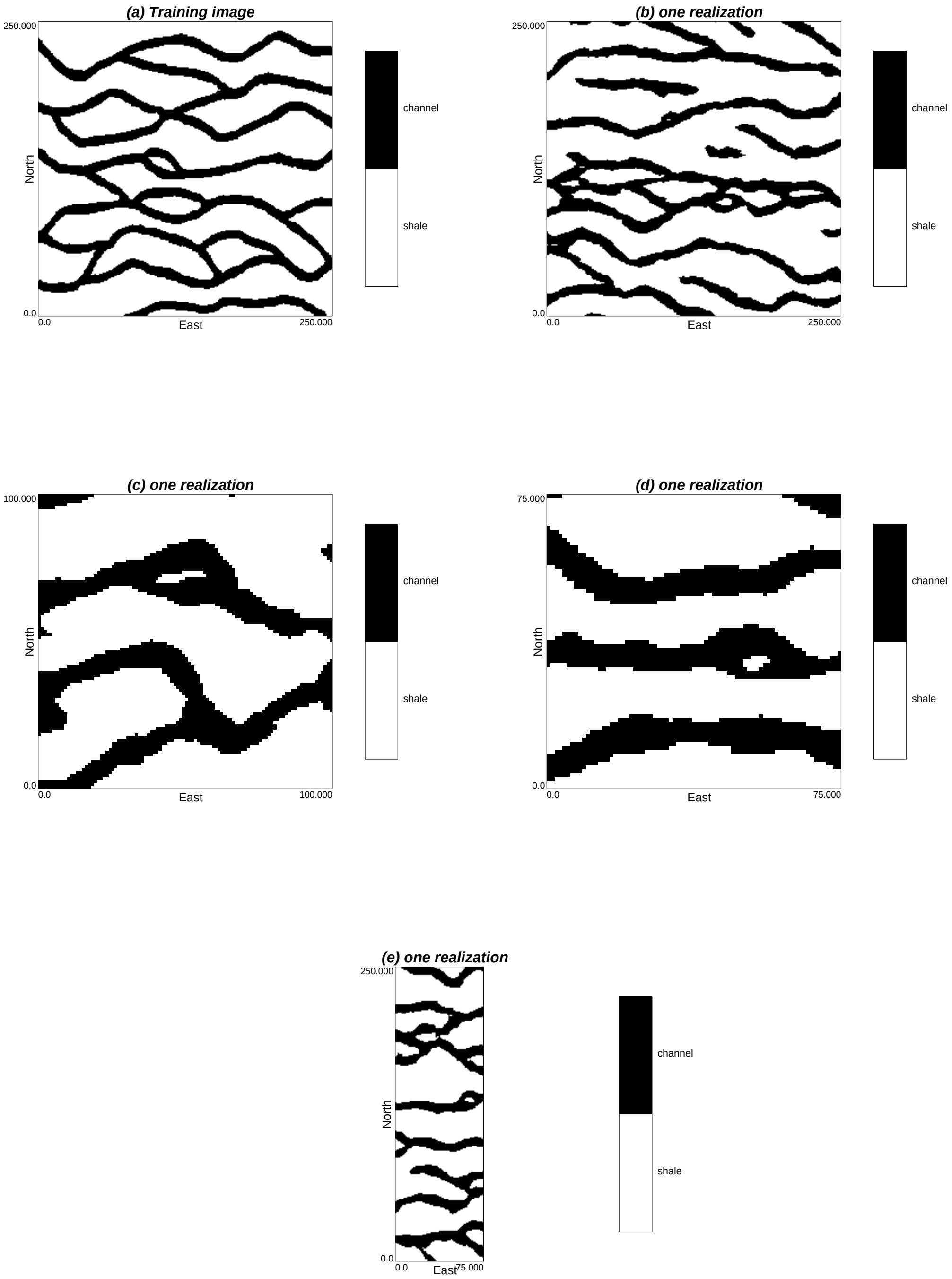

Figure 5: Demonstration of the effect of ergodic fluctuations on snesim, (a) 250×250 training image (b) 250×250 single reservoir model, (c) 100×100 model, (d) 75×275 model, (e) 75×250 model. Note that the each realization has a different scaling of the axis. Compared to other realizations, model (b) has discontinuity in the channel features. This is due to too large ergodic fluctation of the higher-order statistics defined at large distances. Ideally the training image should be more than twice the reservoir size in the direction of largest continuity. This condition is met in cases (c-e), hence no discontinuities are present.

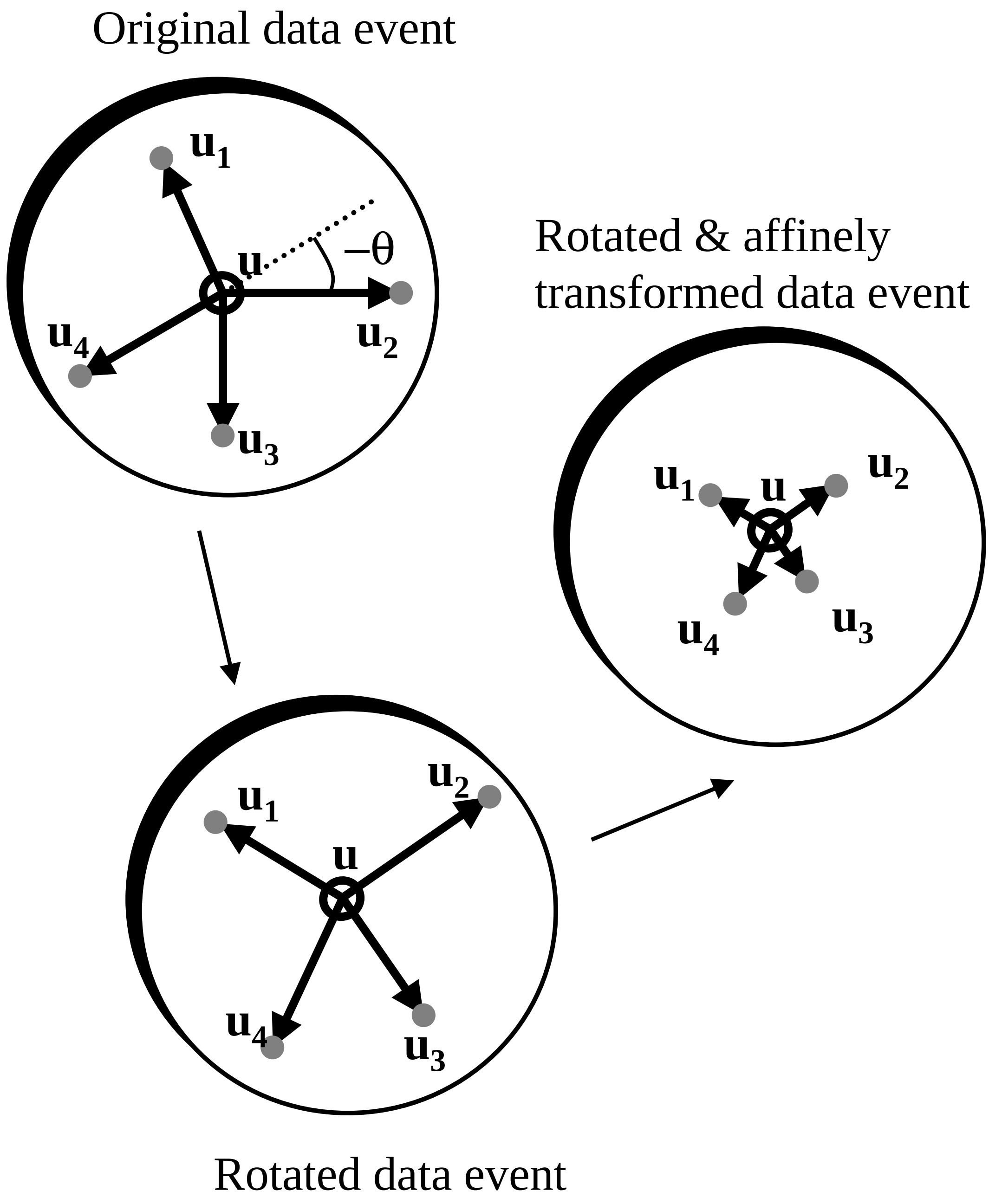

Figure 6: A data event consists of a center location $\mathbf{u}$ plus a set of neighboring well-data and previously simulated node located at coordinates $\mathbf{u}_\alpha$). Each coordinate is first rotated, then affinely transformed (stretching or shrinking).

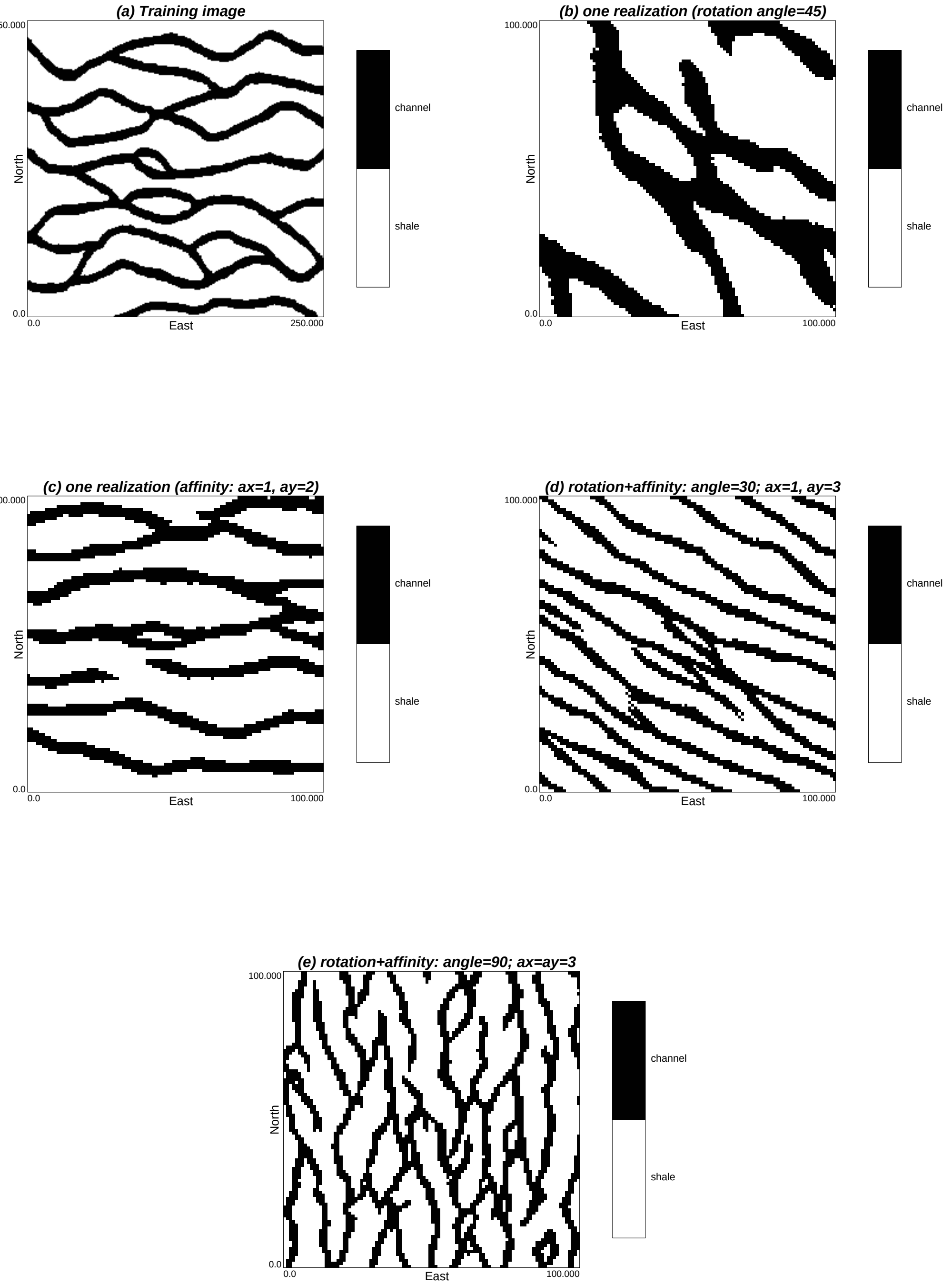

Figure 7: (a) Training image, (b-e) various examples of rotation and affinity transforms. Each realization uses the same training image, displayed in (a).

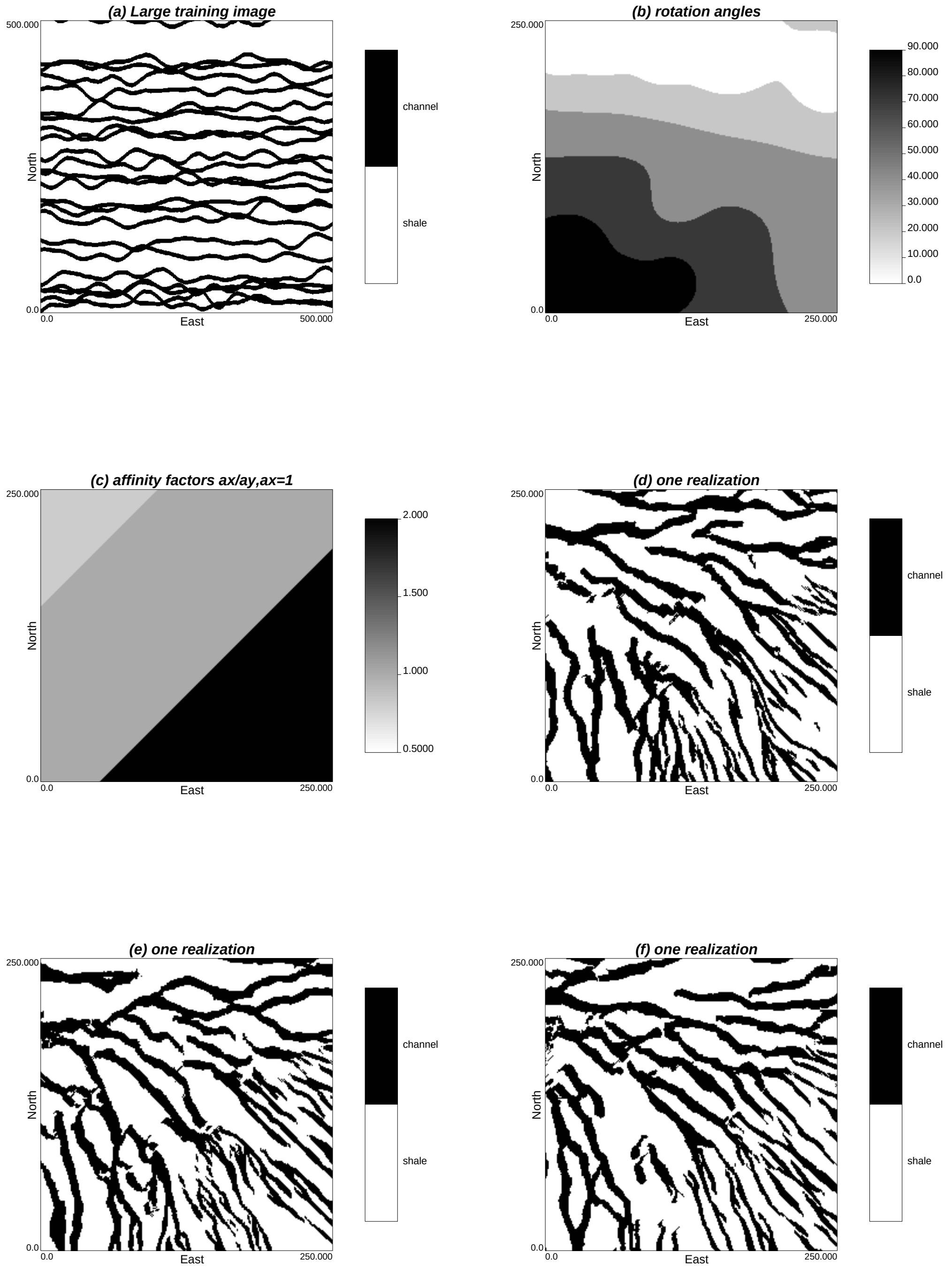

Figure 8: (a) Training image, (b) locally varying rotation angle, (c) locally varying affinity factor a_x/a_y, (d-f) three resulting reservoir models. Using a stationary training image a possibly large variety of reservoir models can be built.

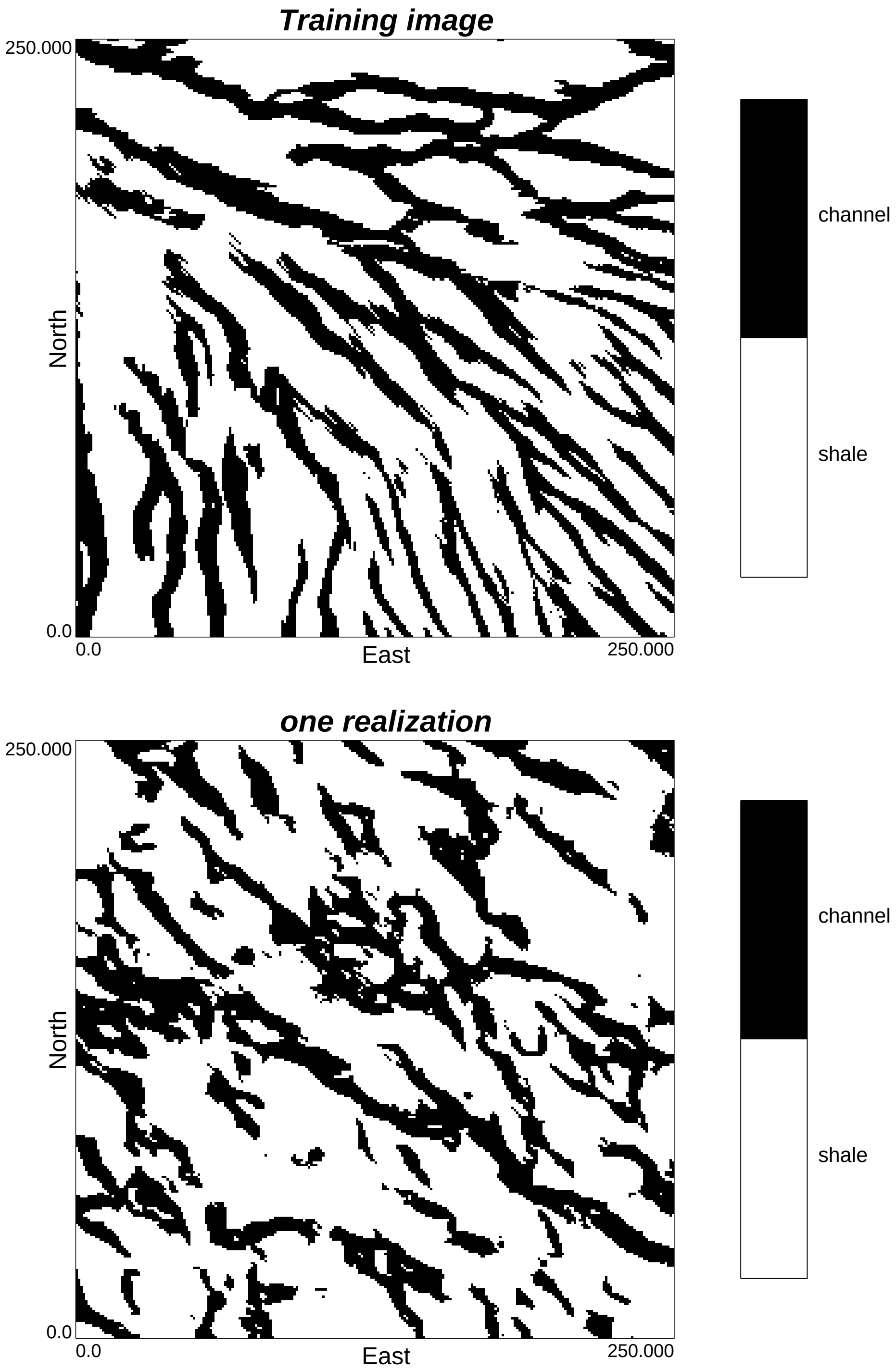

Figure 9: Principle of stationarity: (a) Training image, (b) resulting realization. The use of the non-stationary training image (top) does not provide a reproduction of non-stationary patterns in a model realization (bottom).