

¹ Modeling complex geological structures with elementary training ² images and transform-invariant distances

³ Gregoire Mariethoz¹ and Bryce F. J. Kelly¹

⁴ Received 14 January 2011; revised 21 April 2011; accepted 20 May 2011; published XX Month 2011.

⁵ [1] We present a new framework for multiple-point simulation involving small and
⁶ simple training images. The use of transform-invariant distances (by applying random
⁷ transformations) expands the range of structures available in the simple patterns of the
⁸ training image. The training image is no longer regarded as a global conceptual geological
⁹ model, but rather a basic structural element of the subsurface. Complex geological
¹⁰ structures are obtained whose spatial structure can be parameterized by adjusting the
¹¹ statistics of the random transformations, on the basis of field data or geological context. In
¹² most cases, such parameterization is possible by adjusting two numbers only. This method
¹³ allows us to build models that (1) reproduce shapes corresponding to a desired prior
¹⁴ geological concept and (2) are in phase with different types of field observations such as
¹⁵ orientation, hydrofacies, or geophysical measurements. The main advantage is that the
¹⁶ training images are so simple that they can be easily built even in 3-D. We apply the method
¹⁷ on a synthetic example involving seismic data where the transformation parameters are
¹⁸ data-driven. We also show examples where realistic 2- and 3-D structures are built from
¹⁹ simplistic training images, with transformation parameters inferred using a small number of
²⁰ orientation data.

²¹ **Citation:** Mariethoz, G., and B. F. J. Kelly (2011), Modeling complex geological structures with elementary training images and
²² transform-invariant distances, *Water Resour. Res.*, 47, XXXXXX, doi:10.1029/2011WR010412.

²³ 1. Introduction

²⁴ [2] Numerical subsurface models often rely on scarce
²⁵ data that are insufficient for determining the nature of the geological
²⁶ continuity at a given site. Hence, in most situations
²⁷ the geologist's contextual insights are a crucial source of
²⁸ prior information. In the Earth sciences, a prior model represents
²⁹ this knowledge, i.e., what is known of the geological
³⁰ continuity before the inclusion of location-dependent data.
³¹ In general, the fewer local data available, the greater the importance
³² of the prior model. Prior spatial continuity can be
³³ characterized in various ways. Figure 1 illustrates various
³⁴ prior concepts of geological continuity that can have very
³⁵ different implications in terms of flow and transport properties.
³⁶ They could, for example, correspond to: (1) folded parallel structures of alternating high and low permeability
³⁷ values (Figure 1a and 1b); (2) high permeability values well
³⁸ connected together and defining boundaries between disconnected
³⁹ blocks of low values (Figure 1c and 1d); (3) low
⁴⁰ permeability values well connected together and defining
⁴¹ boundaries between disconnected blocks of high values
⁴² (Figure 1e); and (4) connectivity of high values and connectivity
⁴³ of low values, spatially arranged as elongated bodies
⁴⁴ (Figure 1f).

⁴⁵ [3] Existing geostatistical simulation methods are able
to represent such priors. However, at present only multi-
Gaussian models offer the possibility to easily define the
spatial structure with only a few parameters (mean, variance,
and correlation length). Therefore, the multi-Gaussian framework
offers possibilities for including such parameters in
optimization frameworks [Kitanidis, 1995; Nowak *et al.*,
2010], e.g., finding the variogram parameters that are coherent
with state data such as groundwater head or concentration
measurements.

⁴⁶ [4] Within the multi-Gaussian framework, a wealth of
techniques exist to generate diverse shapes by combining
different Gaussian random fields, such as the plurigaussian
method [Le Loc'h and Galli, 1994; Mariethoz *et al.*, 2009]
or the random coordinate perturbation method and the random
mixture of Gaussian fields [Emery, 2007]. These
methods have the advantage of allowing for parameterization
of the prior through the definition of correlation ranges,
cutoffs, etc. A parametric spatial model is of interest for
model inference and inverse modeling. In certain cases,
however, the assumptions underlying these multi-Gaussian
models do not allow for the reproduction of the behavior of
highly heterogeneous aquifers, especially when connectivity
patterns play an important role [Bastante *et al.*, 2008;
Gómez-Hernández and Wen, 1998; Knudby and Carrera,
2005; Sánchez-Vila *et al.*, 1996; Western *et al.*, 2001].
This can have an especially strong impact when modeling
transport processes [Green *et al.*, 2010; Klise *et al.*, 2009;
Zinn and Harvey, 2003]. Therefore, several alternative simulation
methods have been developed in the past decades.
One approach is to use parametric mathematical models,

¹National Centre for Groundwater Research and Training, University of New South Wales, Sydney, Australia.

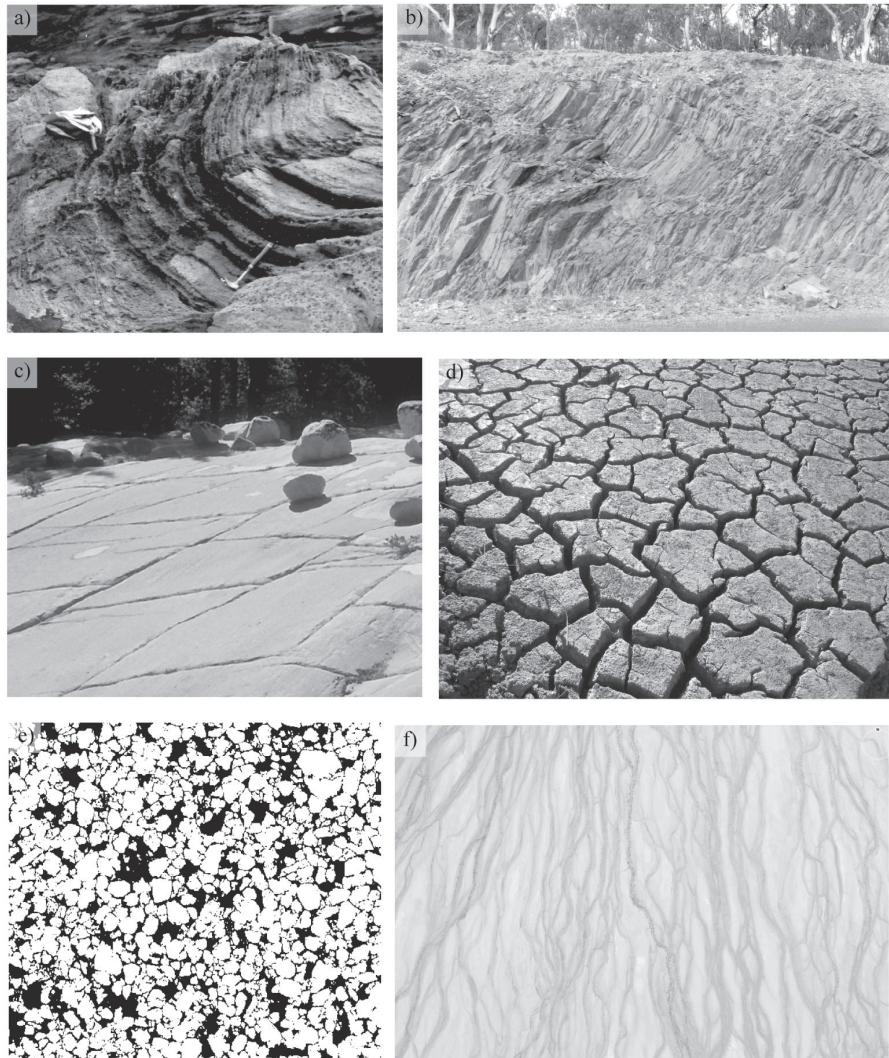


Figure 1. Different prior concepts of geological continuity. (a) Slump folding in mudstone, Warden Head, New South Wales, Australia. (b) Kink folds in interbedded phyllite and quartz greywacke, Kelso-Sofala Road, New South Wales, Australia. (c) Jointed granite, Sierra Mountain, California, U. S. A. (d) Desiccation cracks in Vertosol soil, Narrabri, New South Wales, Australia. (e) Hawkesbury Sandstone thin section. (f) Braided river channels, Diamantina River, Queensland, Australia (image Google Earth, Cnes/Spot, Whereis Sensis, DigitalGlobe). (Photos in panels 1a–1c and e courtesy of Bryce Kelly. Photo in panel 1d courtesy of Anna Greve.)

but to enrich the statistical information by including high-order measures of spatial continuity. Among this promising class of methods spatial cumulants [Dimitrakopoulos *et al.*, 2010] and copulas [Bárdossy and Li, 2008] have recently received a lot of attention.

[5] Other approaches such as object-based [e.g., Deutsch and Wang, 1996; Deutsch and Tran, 2002; Keogh *et al.*, 2007; Pyrcz *et al.*, 2009] and pseudo-genetic models [Cajan *et al.*, 2004; Michael *et al.*, 2010] have been traditionally used to account for the existence of specific connectivity patterns. The main advantage is their ability to define basic shapes that represent geological bodies, using physical representations of the primary geological processes and settings. Conditioning to points data and to non-stationarity in the objects' proportions can be accomplished

through a birth and death algorithm [Allard *et al.*, 2006; Lantuéjoul, 2002], but technical and computational difficulties remain when large amounts of data are present. One of the advantages of the Boolean framework is that the characteristics of the objects used in the simulation can be adjusted to a certain extent (e.g., channel width, sinuosity, etc.).

[6] Recent advances in the domain of geostatistical simulations include multiple-point simulation methods (MPS), which use prior spatial models given by nonparametric conceptual example images, namely training images. The principle of a training image is appealing to geologists because it is a direct representation of a geological concept [Feyen and Caers, 2006]. It has been shown that MPS are able to produce geologically realistic structures by using high-order spatial statistics, which can represent a wide

range of spatially structured, low entropy phenomena, and have been used in a variety of cases [e.g., Caers, 2003; Gonzales et al., 2008; Hu and Chugunova, 2008; Huysmans and Dassargues, 2009; Journel and Zhang, 2006; Lu et al., 2009; Okabe and Blunt, 2007; Ronayne et al., 2008; Wojcik et al., 2009; Wu et al., 2008].

[7] However, although MPS has in recent years become an invaluable tool to integrate geological concepts in subsurface models, the method is still difficult to apply in cases when one does not have enough information to clearly decide which training image to use [de Almeida, 2010; Pyrcz et al., 2008]. Even when the geological context is clear, it is a lengthy geomodeling exercise to build a complex 3-D training image that adequately represents the complexity of geological structures.

[8] Another difficulty is related to the parameterization of the training image. Since the choice of a training image is a discrete decision (either image A or B), it is not straightforward to parameterize it. Suzuki and Caers [2008] proposed a methodology for investigating prior uncertainty in the context of inverse problems through the use of distances between a large variety of discrete scenarios. However, it is still highly desirable to parameterize different multiple-point priors with only a few continuous parameters, in a similar way as multi-Gaussian priors. This is generally not possible because the training images normally used are too complex to be parameterized simply. Training images are generally large (Strebelle [2002] recommends use of a training image much larger than the simulation domain) and ideally contain enough diversity to encompass all possible patterns to be found in the simulation domain.

[9] Even if robust tools exist for multiple-point simulation, the practical difficulties related to building and parameterizing the training image often lead to discarding multiple-point methods and to adopt parametric alternatives for the inferences of spatial continuity, such as traditional variogram-based techniques. Nonetheless, there are cases where more prior information is available than a two-point model of spatial correlation. For example, on a given site, one can often determine very general prior characteristics (e.g., connectivity or disconnectivity of the high values), but the precise types of structures (orientation and dimensions of the structures) are usually uncertain. In these cases, it is desirable to have methods that integrate such prior characteristics in stochastic models while leaving unknown characteristics random.

[10] In this paper we present a new concept and algorithm for multiple-point simulation that considers elementary/simplistic training images that are an expression of prior spatial continuity. During the simulation the diversity of patterns is enriched by applying random transformations. It is equivalent to comparing patterns up to a transformation, or using a transform-invariant distance. This method enables the generation of complex geological images whose spatial structure can be parameterized by adjusting the statistics of the random transformations. In most cases, such parameterization is possible by adjusting only two numbers. It allows parameterizing multiple-point priors in a straightforward and integrated manner. Moreover, it also resolves the problem of building a training image because building elementary images is easy even in 3-D. The training image is no longer regarded as a global conceptual

geological model, but rather a basic structural element of the subsurface, such as the shapes used in object-based simulations. These basic elements can be parameterized to create models that (1) reproduce shapes corresponding to a desired prior geological concept, and (2) are in phase with different types of field observations such as orientation, hydrofacies, or geophysical measurements.

[11] Section 2 of the paper is a short description of the direct sampling method, which is the geostatistical simulation method used to illustrate our concept. Note, however, that the fundamental idea could, in theory and regardless of implementation issues, be applied to any simulation algorithm that is based on the computation of a distance between patterns. Section 3 details the concept of transform-invariant distances and elementary training images, and illustrates this concept with basic examples. Section 4 shows how conditioning data can be used to structurally guide the simulation. Finally, Section 5 shows how to introduce nonstationarity in the transformation parameters to model complex geological structures.

2. Background on the Direct Sampling Method

[12] The direct sampling method is described in details in the work of Mariethoz et al. [2010] as a multiple-point simulation tool, although it was found that a similar algorithm had been proposed a decade earlier by Erfos and Leung [1999] in the context of texture synthesis. It generates conditional multiple-point simulations that present the same spatial dependence as a user-defined training image. Note that we use the term spatial dependence and not spatial correlation to emphasize that no linearity is assumed. Direct sampling is a sequential simulation method. It proceeds by successively visiting all locations \mathbf{x} of a regular grid, each time assigning a value $Z(\mathbf{x})$ that depends on the values of the neighboring locations of \mathbf{x} , denoted $Z(\mathbf{x}_1)$, $Z(\mathbf{x}_2)$, ..., $Z(\mathbf{x}_n)$, and on the lag vectors defining the position of each neighbor relative to \mathbf{x} , denoted \mathbf{h}_1 , \mathbf{h}_2 , ..., \mathbf{h}_n . For simplicity, we denote N_x the neighborhood of \mathbf{x} , comprising both the lag vectors and the values at the neighboring locations. Figure 2 illustrates the different components of N_x .

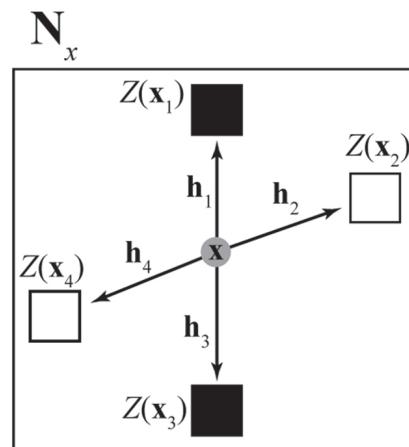


Figure 2. The different elements defining N_x , the neighborhood of the location x .

[13] The direct sampling algorithm is summarized below:

[14] Input: (1) Simulation grid SG with locations denoted \mathbf{x} , (2) training image TI with locations denoted \mathbf{y} , (3) distance function $d(\cdot)$ bounded in $[0,1]$, (4) distance threshold t (distance under which two neighborhoods are considered identical).

[15] 1. **While** # noninformed locations in SG ≥ 0 **do**.

[16] 2. Choose a noninformed location \mathbf{x} in SG and identify the neighborhood \mathbf{N}_x centered on \mathbf{x} .

[17] 3. Initialize $d(\mathbf{N}_x, \mathbf{N}_y) = \infty$.

[18] 4. **While** $d(\mathbf{N}_x, \mathbf{N}_y) \geq t$ **do**.

[19] 5. Sample a location \mathbf{y} in TI.

[20] 6. Define \mathbf{N}_y as the neighborhood centered on \mathbf{y} having identical lag vectors as \mathbf{N}_x .

[21] 7. Compute the distance $d(\mathbf{N}_x, \mathbf{N}_y)$.

[22] 8. **End While**.

[23] 9. Assign $Z(\mathbf{x}) = Z(\mathbf{y})$.

[24] 10. **End While**.

[25] Output: Completed simulation SG.

[26] The main principle is that as soon as a matching neighborhood is found (i.e., when $d(\mathbf{N}_x, \mathbf{N}_y) < t$, one can consider that $(\mathbf{N}_y \equiv \mathbf{N}_x)$, the corresponding value $Z(\mathbf{y})$ is a sample of the conditional distribution Prob $\{Z(\mathbf{x})|\mathbf{N}_x\}$.

[27] The distance $d(\mathbf{N}_x, \mathbf{N}_y)$ can be computed in different ways depending on the nature of the variable Z . It was shown by Mariethoz et al. [2010] that choosing the appropriate distance allows one to adapt the method to either categorical or continuous variable cases, and also to jointly simulate several dependant variables. The distances proposed include, for a categorical variable,

$$d(\mathbf{N}_x, \mathbf{N}_y) = \frac{1}{n} \sum_{i=1}^n a_i \quad \in [0, 1], \quad (1)$$

where $a_i = \begin{cases} 0 & \text{if } Z(\mathbf{x}_i) = Z(\mathbf{y}_i) \\ 1 & \text{if } Z(\mathbf{x}_i) \neq Z(\mathbf{y}_i) \end{cases}$,

and for a continuous variable, the normalized Manhattan distance

$$d(\mathbf{N}_x, \mathbf{N}_y) = \frac{1}{n} \sum_{i=1}^n \frac{|Z(\mathbf{x}_i) - Z(\mathbf{y}_i)|}{\max_{y \in TI} Z(y) - \min_{y \in TI} Z(y)}. \quad (2)$$

[28] Since the distances are usually defined such that they are within the interval $[0,1]$, the threshold is also bound to the same interval. Defining a threshold of $t = 0$ means that the patterns of the training image will be reproduced with the highest possible accuracy. Conversely, when setting $t = 1$, the algorithm unconditionally samples values from the training image, therefore reproducing the marginal distribution of Z but with no spatial dependence. Between these two extreme cases, the value of t determines how accurately the patterns of the training image are reproduced.

3. Transform-Invariant Distances

3.1. The Principle of Transform-Invariant Distances

[29] Let us define a geometrical transformation T_p applied to \mathbf{N}_x , where p is a parameter affecting the transformation.

Several types of geometrical transformations can be considered. For now, we consider two types of transformations, which are (1) rotations centered on \mathbf{x} , denoted $R(\cdot)$, and (2) affinity (or homothetic) transforms, denoted $A(\cdot)$. Note that we consider only geometrical transforms that affect the lag vectors \mathbf{h} while the values $Z(\mathbf{x}_i)$ remain unchanged. For rotations, the parameter p corresponds to the rotation angle and for affinity transforms, it corresponds to the affinity ratio. Figure 3 shows a few examples of transformations applied to a simple neighborhood (or pattern) \mathbf{N}_x .

[30] The direct sampling method uses a distance between a neighborhood \mathbf{N}_x centered on the location \mathbf{x} , and another neighborhood \mathbf{N}_y centered on location \mathbf{y} in the training image (see step 7 of the algorithm). Instead, we propose a consideration of the distance between the ensemble of possible neighborhoods resulting from geometric transforms of \mathbf{N}_x and a neighborhood \mathbf{N}_y in the training image. By doing this, we compare the neighborhoods \mathbf{N}_x and \mathbf{N}_y up to a transformation [i.e., there exists a transformation $T(\cdot)$ that makes $T(\mathbf{N}_x)$ identical to \mathbf{N}_y]. When considering transform-invariant distances, Figure 3a, 3b, and 3c are identical up to a rotation, and Figure 3a, 3d, and 3e are identical up to an affinity. We define a transform-invariant distance as

$$d^T(\mathbf{N}_x, \mathbf{N}_y) = \operatorname{argmin}\{d(T_p(\mathbf{N}_x), \mathbf{N}_y)\} : \forall p \text{ in } [a,b], \quad (3)$$

where a and b are bounds specified for parameter p . A transform-invariant distance can be computed by finding a value of p (in the interval $[a,b]$) such that the distance $d(T_p(\mathbf{N}_x), \mathbf{N}_y)$ is minimal. Once the value of p is found, the corresponding distance is the transform-invariant distance. Note that the parameter p is not prescribed in advance, but determined by a search procedure that evaluates all p values in $[a,b]$ and finds one corresponding to an acceptable distance (or mismatch) between both neighborhoods.

[31] Since the direct sampling method only needs to sample a single neighborhood \mathbf{N}_y whose distance to \mathbf{N}_x is lower than the threshold t , using a transform-invariant distance simply consists of applying to \mathbf{N}_x a transformation with a random p -value in $[a,b]$ each time the distance $d(\mathbf{N}_x, \mathbf{N}_y)$ is computed. This is accomplished by applying the random transformation at step 7 of the direct sampling algorithm presented in Section 2. As soon as $d(T_p(\mathbf{N}_x), \mathbf{N}_y) \leq t$, one can consider that $\mathbf{N}_y \equiv \mathbf{N}_x$ (up to a transformation T_p) and assign $Z(\mathbf{x}) = Z(\mathbf{y})$. This sampling procedure is equivalent to searching the entire interval $[a,b]$, but much less computationally demanding.

[32] When using transform-invariant distances in multiple-point simulations, one does not only consider the patterns found in the training image, but all their possible transformations in the range $[a,b]$. This extra degree of freedom constitutes a considerable increase in the diversity of the possible structures that can be produced, and allows complex spatial patterns to be generated, even from simple training images. By adjusting the values of the transformation parameter bounds (either constant or spatially variable), it is possible to parameterize the ensemble of patterns available for the simulation. The consequence is a controlled increase in the variability between realizations. Emery and Ortiz [2011] recently showed that a robust use of a complex training image would call for a dramatic increase in the training image size, to the point that it

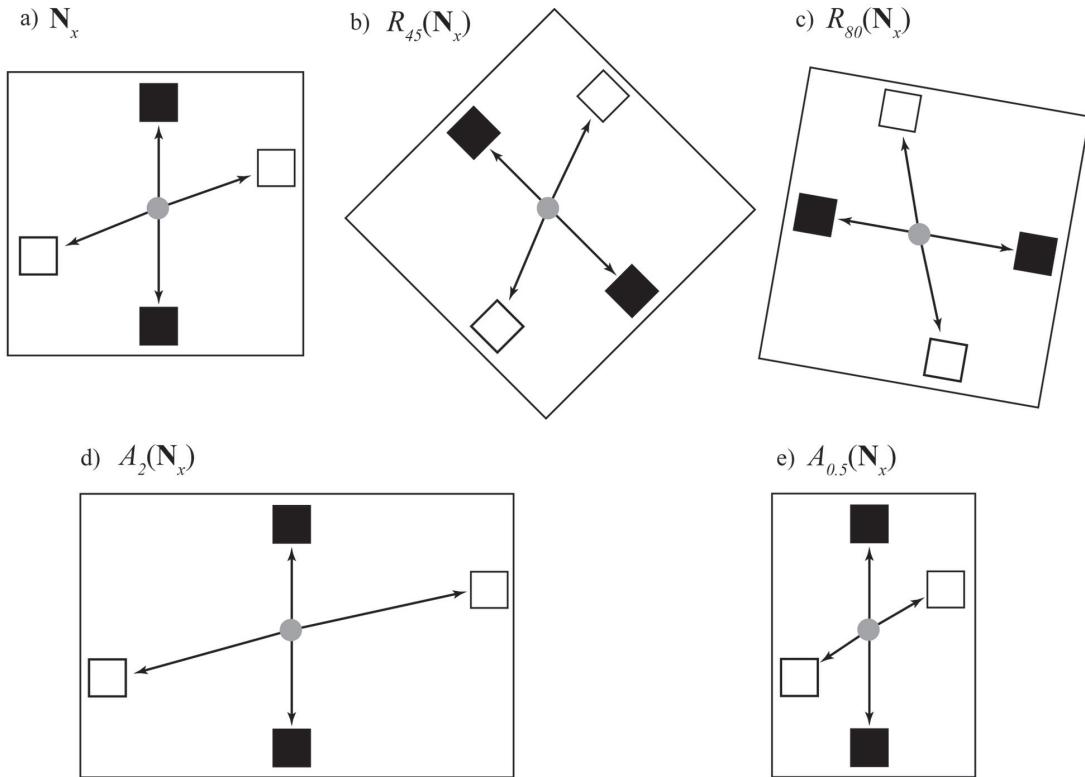


Figure 3. Different transformation applied to N_x . (a) Original neighborhood N_x . (b) Forty-five-degree rotation. (c) Eighty-degree rotation. (d) Factor 2 affinity along the X axis. (e) Factor 0.5 affinity along the X axis.

would be computationally unfeasible. It was often observed that training images contain certain idiosyncrasies that are not repetitive enough and therefore appear unchanged in the simulations. A good example is the “islands” in the classical channels training image in the work of Strebelle [2002]. In several cases one can find simulations where these islands are reproduced almost identically as in the training image. This lack of variability cannot appear with our method because the transform-invariant distances produce a large number of patterns from a single idiosyncrasy, therefore increasing the robustness of the multiple-point statistics.

[33] The distance in equation (3) could in theory be used with all simulation algorithms that rely on a distance between patterns, such as Simpat [Arpat and Caers, 2007], Filtersim [Zhang et al., 2006], or Dispat [Honarkhah and Caers, 2010]. In particular, the Dispat method would be very well suited for this purpose since it also uses a distance to expand the space of the patterns that exist in the training image. However, these methods store all the patterns used for the simulation, and the increased pattern diversity brought by transform-invariant distances may lead to significant memory requirements and potential computational issues. In comparison, direct sampling does not have any storage constraints.

3.2. Transform-Invariant Distances Put in Practice

[34] To illustrate the method, we use the simplest possible training image that consists of a binary variable depicting parallel horizontal lines. This training image is

simplistic, but represents a geologically meaningful spatial continuity. Concepts such as global connectivity or disconnection can be conveyed even at the scale of an elementary training image, which is much smaller than the simulation domain. We illustrate this by producing various spatial patterns with different transform-invariant distances. The elementary training image is displayed in Figure 4a, and realizations corresponding to different transformation parameters are shown in Figure 4b–4j. For all realizations in Figure 4, the direct sampling method was used with a distance threshold set to 0 and a neighborhood consisting of the closest 25 neighbors. The size of the training image is 100×100 pixels and the realizations are 500×500 pixels. Note that using such a small training image does not adhere to the general principle that recommends using a training image larger than the simulations. Using a larger version of this training image with standard multiple-point methods would not enrich the available patterns because those are repetitive. However, the realizations of Figure 4 show that using transform-invariant distances greatly increases the diversity of patterns.

[35] Figure 4b shows a realization that does not use transformations (i.e., rotations of 0° and an affinity factor of 1). It corresponds to what one would expect from a traditional multiple-point simulation method, with parallel lines that are sometimes disconnected because the training image does not inform the large-scale connectivity. Figure 4c is identical, except that a deterministic rotation of 45° is uniformly applied to all of the patterns. The result is similar to

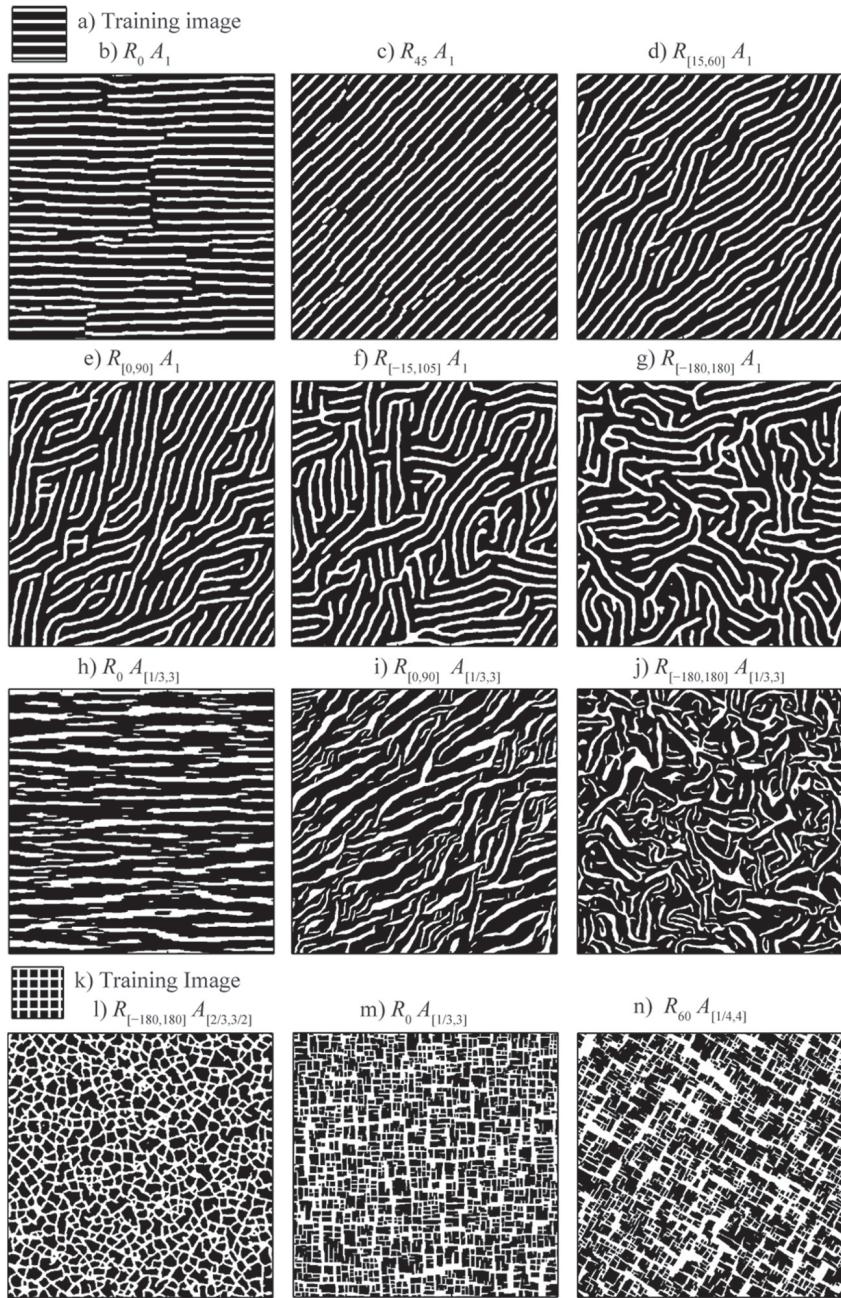


Figure 4. Examples of textures obtained with transform-invariant distances. (a) Lines training image. (b–j) corresponding realizations with various transformation parameters. (k) Checkerboard training image. (l–n) corresponding realizations with various transformation parameters. See the text for a detailed discussion.

that obtained with the transformations method proposed by Strebelle and Zhang [2004]. Unlike our approach, this method does not allow for an increase in the number of patterns available for simulation, neither does it allow for accounting the uncertainty in the transformation parameters. Additionally, it increases the computational load by creating as many data event catalogues as the number of categories of transformation parameters.

[36] Figure 4d is different because here the rotations are randomly distributed between 15° and 60° . The transform-

invariant distance allows assembling patterns that respect the general continuity of the training image, even if those patterns are not oriented in the same direction. In Figure 4d, this produces coherent lines, but with a small degree of sinuosity. Since the patterns of the training image are reproduced up to a rotation, the pool of acceptable patterns is larger, resulting in increased continuity of the simulated lines. The diversity of patterns is not contained in the training image itself, but in the large amounts of possible transformations. Note that the lines almost never intersect

395 because patterns of intersection do not appear in the training
 396 image. Figure 4e uses angles between 0° and 90° , and it
 397 is visible that the extreme orientations of the structures
 398 tend to be either horizontal or vertical. Figure 4f shows
 399 even more variability in the orientations, and Figure 4g,
 400 with angles between -180° and 180° allows a total free-
 401 dom for the orientations, resulting in maze-like structures.
 402 Figure 4h–4j include, in addition to rotations, random affin-
 403 ity transforms ranging between 1/3 and 3. When rotations
 404 and affinities are used simultaneously, the added diversity
 405 of patterns creates shapes that can resemble geological
 406 structures such as clay drapes seen in cross-section (Figure
 407 4h) or anastomosed channels seen in plan view (Figure 4i).

[37] In these examples, the training image only carries a
 409 concept of “elongated and connected bodies” that can ma-
 410 terialize in various ways under the different transfor-
 411 mations applied. Another case with a different training image is
 412 also represented in Figure 4k–4n. Figure 4k shows a train-
 413 ing image carrying a concept corresponding to “discon-
 414 nected bodies separated by interconnected interfaces.”
 415 Using random rotations between -180° and 180° , it is pos-
 416 sible to obtain structures resembling desiccation cracks or,
 417 if applied in 3-D, certain types of porous media (Figure 4l).
 418 With random affinity transforms, the results tend to look
 419 like fractured or karstic media observed in areas with or-
 420 thogonal stress fields or regions of overburden stress relief
 421 (Figure 4m and 4n). Note that this exercise could be contin-
 422 ued using different elementary training images and trans-
 423 formation parameters to generate various families of
 424 structures corresponding to the cases presented in Figure 1.

[38] Although the transformations are randomly applied,
 426 spatial coherence is preserved throughout all of the exam-
 427 ples. This is possible because the transformation parame-
 428 ters at each location are determined on the fly when
 429 simulating each value $Z(x)$. For each node, the formulation
 430 of equation (3) implies that the only acceptable transforma-
 431 tions are the ones that yield patterns compatible with the
 432 training image. Therefore, it is ensured that the local trans-
 433 formations are consistent with the prescribed geological
 434 continuity. If the transformations were determined in
 435 advance, for example, using a map with an assigned rota-
 436 tion value at each location, the compatibility between the
 437 transformation parameters and the training image patterns
 438 would not be guaranteed.

439 4. Data-Driven Transformations

[39] A traditional conceptualization of geostatistics is
 441 that the spatial model is first derived from data, and then
 442 used for estimation or simulation. On the other hand, multi-
 443 ple-point simulation considers cases where the local data
 444 alone are not sufficient to infer a spatial model. However,
 445 conditioning point data, especially if present in large
 446 amounts, can influence the simulated structures and hence
 447 the underlying random function.

[40] As noted by Mariethoz and Renard [2010] one can
 449 distinguish two kinds of constraints that apply to determin-
 450 ing the value of each pixel: the structural constraints and
 451 the local constraints. Structural constraints are imposed by
 452 the spatial model, which in this case consists of patterns of
 453 the training image, considered up to a transformation. To
 454 Local constraints are given by conditioning point data. To

obtain consistent simulations, these two types of constraints
 455 must be compatible. An example of incompatibility is the
 456 use of conditioning data showing continuity along the X
 457 axis with a training image that depicts channels running
 458 along the Y axis.

[41] When using transform-invariant-distances, the struc-
 460 tural model is flexible and can adapt itself to the condition-
 461 ing data. The minimization argument in equation (3) means
 462 that the transformation parameters are always adapted to
 463 maximize compatibility with the neighboring pixels, with
 464 respect to both structural and local constraints. Therefore,
 465 when conditioning point data are present, the issue of
 466 incompatibility between training image patterns and condi-
 467 tioning data is not as acute as with the classical approach.
 468 The patterns are automatically adapted to obtain as much
 469 compatibility as possible, and the final result is that the
 470 simulated structures are seamlessly wrapping around the
 471 data without user intervention.

[42] A practical example would be when integrating
 473 interpreted structural surfaces in a geological model (which
 474 could have been derived from a geophysical survey, for
 475 example, a seismic profile). This information is often char-
 476 acterized by a series of points along known interfaces
 477 (Figure 5a). The subsurface models should present struc-
 478 tures globally oriented along the delineated interfaces, but
 479 showing some variability away from these known interfaces.
 480 This problem is easily dealt with using rotation-invariant
 481 distances and an elementary training image (here we use
 482 the lines training image shown in Figure 5b). Even if the
 483 data show relatively complex orientations, the patterns
 484 within the training image are oriented to maximize the co-
 485 herence with the conditioning point data. As a conse-
 486 quence, the realizations (Figure 5c) respect the structures
 487 within the training image (structural constraints), but with
 488 orientations that are compatible with the data (local con-
 489 straints). The distance threshold is set to 0 and the number
 490 of neighbors to 25. Figure 5d shows the probability of occur-
 491 rence of the white facies (conductive layers) computed after
 492 100 realizations. In the vicinity of the delineated interfaces,
 493 strong local constraints result in very well-defined orienta-
 494 tions, with a series of three or four parallel layers that are
 495 present in most realizations. Away from the data, less local
 496 constraints are present and more variability between realiza-
 497 tions is observed in the orientation of the layers.

5. Using Nonstationary Transformation 499 Parameters for Structural Modeling

[43] Nonstationarity is often defined as a parametric
 501 trend in the values of the variable considered (e.g., a poly-
 502 nomial function). Such trends can be modeled in a variety
 503 of ways [see Goovaerts, 1997]. On the other hand, struc-
 504 tural nonstationarity (such as a transition from channels to
 505 lobes) is usually too complex to be modeled with a few pa-
 506 rameters and needs to be defined in a nonparametric way.
 507 Several authors have proposed methods of modeling non-
 508 stationarity in the context of multiple-point simulation
 509 [Chuganova and Hu, 2006; De Vries et al., 2009; Straub-
 510 haar et al., 2011], but these methods require either very
 511 complex training images that are difficult to obtain, or a
 512 very precise zonation of the domain that can be quite cum-
 513 bersome in practice.

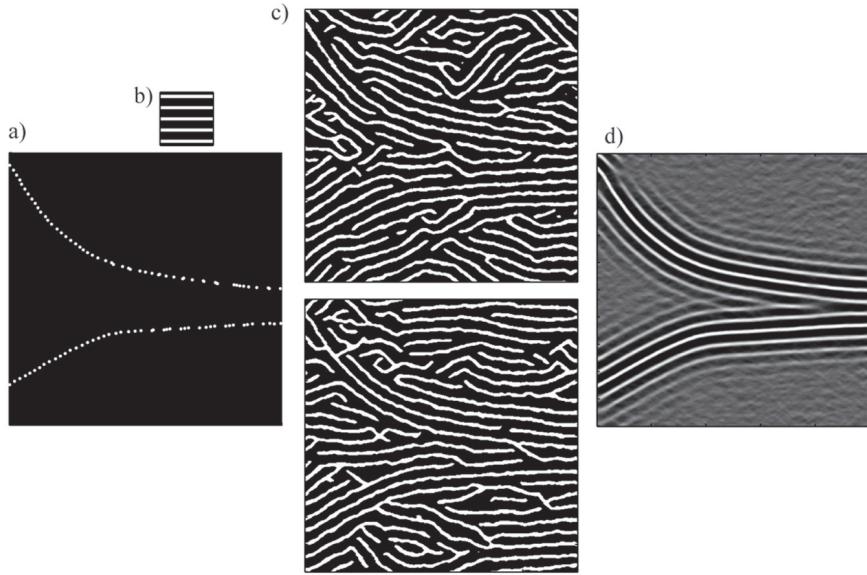


Figure 5. Illustration of the integration of interfaces derived from seismic profiles. (a) Pointed locations along interfaces (derived from seismic profile). (b) Elementary training image. (c) Two different data-driven realizations. (d) Probability of conductive layer (white) computed using 100 realizations.

[44] The examples shown in Figure 4 use homogeneous transformation parameters (i.e., identical bounds for transformation parameters are applied at all locations of the domain). In this section we propose to model structural nonstationarity simply by imposing nonstationary transformation parameters on the simulation domain.

5.1. Modeling Structural Folding With Rotation Operations

[45] The focus of this section is to consider variable transformation parameters to model nonstationary geological structures, especially the ones involving rotations such as folds. Such geological modeling is usually accomplished using geomodeling techniques [Mallet, 2008]. Among these, the implicit approach, based on level set methods, has received a lot of interest in the last few years [Lajaunie *et al.*, 1997; Maxelon and Mancktelow, 2005; Maxelon *et al.*, 2009]. Its main principle is to define 3-D scalar fields (or potential fields) whose isosurfaces are the geological interfaces. The potential fields are obtained by interpolation of local facies and orientation data. While facies measurements give the value of the potential at specific locations, orientation data provide information on the gradient of the potential field, and can thus be used to guide the interpolation. One advantage of these approaches is that they allow one to model a set of surfaces accounting for all orientation data with a single interpolation. In fact, a 3-D model can be built using only the simplest field measurements taken by geologists consisting of facies and orientation data, namely strike and dip measurements. In this section we show how our approach can be used as an alternative for obtaining 3-D models based on similar field measurements.

[46] Let us consider M angle measurements m_i , $i = 1 \dots M$. For simplicity, we only consider the dip angle (i.e., the angle between the horizontal plane and the inclined surface), but the approach is general and can be applied to

strike and dip together. Interpolation is needed in order to obtain the transformation parameters (orientations) on the entire domain. To this end, we will see that it is more convenient to express the bounds $[a, b]$ for the rotations as a median value $\alpha(\mathbf{x})$ and a tolerance $\varepsilon(\mathbf{x})$. For example, using $\alpha = 30$ and $\varepsilon = 10$ is equivalent to using the bounds $[20, 40]$.

[47] At the location \mathbf{x}_i of an orientation measurement, assuming no measurement error, one can safely define $\alpha(\mathbf{x}_i) = m_i$ and $\varepsilon(\mathbf{x}_i) = 0$. This is equivalent to using infinitely narrow bounds in the transformation parameters: $[a, b] = [m_i, m_i]$. At locations further away from the measurement, the bounds defining the orientation should be enlarged and become $[a, b] = [\alpha(\mathbf{x}) - \varepsilon(\mathbf{x}), \alpha(\mathbf{x}) + \varepsilon(\mathbf{x})]$. We propose the use of simple interpolation techniques to obtain the transformation parameters $\alpha(\mathbf{x})$ and $\varepsilon(\mathbf{x})$ on the entire simulation domain. The most logical choice would be to use kriging for this interpolation. A major advantage would then be that the tolerance ε could be directly related to the estimation variance. The final result would provide an estimation of the angle and the tolerance on the entire domain that would be consistent with the field measurements. Unfortunately, in most practical cases the orientations are nonstationary and the measurements too few to properly compute and adjust a variogram. Hence, we propose a simple inverse square distance interpolation of the angles. The interpolated angles are then

$$\alpha(\mathbf{x}) = \sum_i \frac{w_i(\mathbf{x})\alpha_i}{\sum_j w_j(\mathbf{x})}, \quad w_i = \frac{1}{\|\mathbf{x} - \mathbf{x}_i\|^2}. \quad (4)$$

[48] Note that the angles actually represent a continuum with the highest values being similar to the lowest. For example, angles of 175° and -175° are only 10° apart, hence particular care has to be taken to avoid considering them as 350° apart. To take account for this, we use an

583 intermediate step proposed by *Gumiaux et al.* [2003],
 584 which addresses the issue by separately interpolating the
 585 sine and the cosine of α . Here we use equation (4) for this
 586 interpolation step. Reconstruction of $\alpha(\mathbf{x})$ is then done over
 587 the entire domain using its sine and cosine values.

588 [49] The tolerance is defined in an ad hoc way by consider-
 589 ing the ratio between the distance to the closest data loca-
 590 tion and the domain size S , and multiplying it by a
 591 maximum tolerance ε_{\max} :

$$\varepsilon(\mathbf{x}) = \frac{\arg \min(\|\mathbf{x} - \mathbf{x}_i\|)}{S} \varepsilon_{\max}. \quad (5)$$

592 [50] The goal is to have $\varepsilon(\mathbf{x}) = 0$ at the orientation mea-
 593 surement locations and a maximum tolerance of ε_{\max} (i.e.,
 594 the highest uncertainty) at the locations that are the furthest
 595 away from the measurements. We use such basic interpola-
 596 tion of the rotation parameters because we consider cases
 597 where only a handful of orientation measurements are
 598 available. A variety of angle interpolation techniques could
 599 be used for this. For a comprehensive review of interpola-
 600 tion methods applied to angles the reader can refer to
 601 *Gumiaux et al.* [2003]. Once $\alpha(\mathbf{x})$ and $\varepsilon(\mathbf{x})$ are known over
 602 the entire simulation domain, realizations can be generated
 603 as previously with distance (1), the only difference being
 604 that the transformation parameters are spatially variable.

605 [51] Figure 6 illustrates the proposed methodology.
 606 Three orientation measurements are displayed by vectors
 607 on Figure 6a. Figure 6b and 6c represent, respectively, the
 608 result of the interpolated angles (in degrees) using equation
 609 (4), and the tolerance obtained using equation (5) with a
 610 maximum tolerance ε_{\max} of 90°.

611 [52] For the simulation, two different training images are
 612 considered (Figure 6d). The first one is continuous and rep-
 613 presents hydraulic conductivity. It consists of one unconditional

615 multi-Gaussian realization of 200×200 pixels generated
 616 by the spectral method [*Le Ravalec-Dupin et al.*, 2000]. An
 617 exponential variogram model is used with a mean of −4
 618 and a variance of 2. A strong anisotropy is represented by
 619 ranges of 120 in the X direction and 5 in the Y direction.
 620 The second training image, consisting of horizontal lines
 621 representing hydrofacies, is the same as in Figure 4a. One
 622 realization is generated with each training image. The dis-
 623 tance threshold is set to 0 for the categorical simulation,
 624 0.02 for the continuous simulation, and the number of
 625 neighbors is set to 25 in both cases. These two realiza-
 626 tions are shown in Figure 6e and 6f. For both training images,
 627 the structures of the training image are oriented in a way
 628 that is consistent with the initial measured orientations
 629 $m(\mathbf{x}_i)$. Moreover, the continuities are well preserved, with
 630 elongated bodies of high hydraulic conductivity in the con-
 631 tinuous case and parallel channels of constant thickness in
 632 the categorical case.

633 [53] To investigate the sensitivity of the maximum toler-
 634 ance ε_{\max} , we perform exactly the same exercise with
 635 $\varepsilon_{\max} = 180^\circ$. As a result, the interpolated angles remain the
 636 same but all tolerance values are doubled compared to
 637 Figure 6c, globally increasing the variability of orienta-
 638 tions. The resulting simulations, shown in Figure 7, indeed
 639 present more erratic channel orientations while being con-
 640 sistent at the measurement locations.

641 [54] To compare the channel orientations over an ensem-
 642 ble of realizations, we perform a skeleton analysis of 100
 643 categorical realizations with (1) $\varepsilon_{\max} = 90^\circ$ and (2)
 644 $\varepsilon_{\max} = 180^\circ$. Skeleton analysis is a morphological opera-
 645 tion that involves removing pixels on the boundaries of
 646 objects, but not allowing the objects to break apart. As a
 647 result, each channel is represented by a line that runs
 648 through its center. The results are shown in Figure 8. Note
 649 that in order to keep the figure readable, we simultaneously

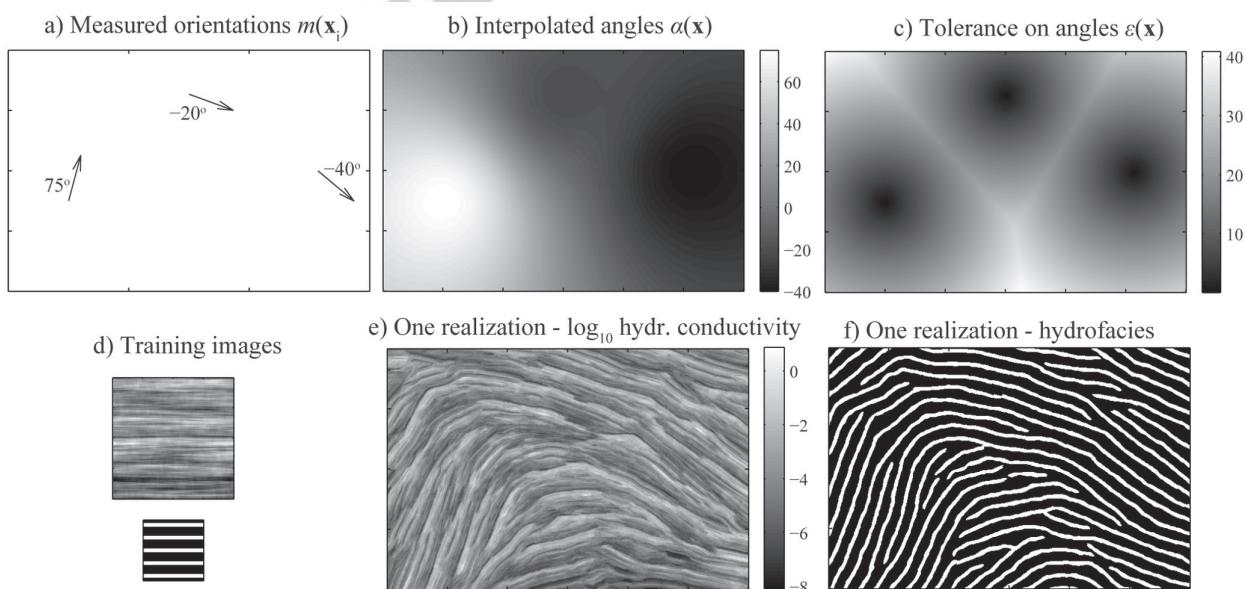


Figure 6. Application of nonstationary transformation parameters to model a fold in 2-D, with $\varepsilon_{\max} = 90^\circ$. (a) Orientation measurements. (b–c) Interpolated rotation parameters. (d) Two different training images. (e–f) corresponding realizations.

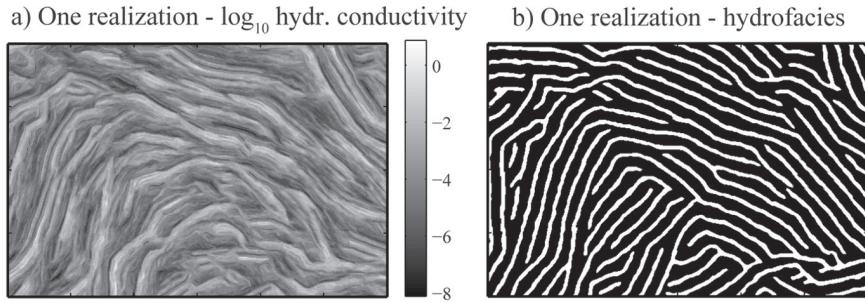


Figure 7. Identical example with a value of $\varepsilon_{\max} = 180^\circ$. (a) Continuous variable. (b) Categorical variable.

represent the skeleton analysis of only five realizations in each case. In both cases, the channels are all correctly oriented at the data locations (thus, resulting in parallel lines), but they can represent a large amount of variability at locations distant from the data, such as the lower central portion or the top left corner of the image, where the lines cross each other, showing different orientations from one realization to the next. The orientation variability is much lower in Figure 8a than in 8b due to the lower ε_{\max} value that controls the tolerance on angles. However, within a single realization the channels remain relatively parallel to each other, as shown in Figure 7b, due to the parallel structures found in the training image.

[55] Increasing the number of orientation measurements allows obtaining geological structures of increasing complexity. Figure 9 illustrates a situation with eight orientation measurements, with the same training images as used in Figure 6. The interpolated angles result in a succession of two different folds. Figure 10 shows an example where even more complexity has been introduced using distances that are both rotation- and affinity-invariant. In addition to the rotation and tolerances shown in Figure 10b and 10c, we include random affinity transforms ranging uniformly between 1/3 and 3 on the entire domain. The resulting structures have variable widths and are not parallel any more, corresponding to sedimentary structures observed in deltaic environments. Note that both examples of Figures 9 and 10 use a value of $\varepsilon_{\max} = 90^\circ$.

5.2. High-Resolution 3-D Example

[56] Within the framework of classical multiple-point simulation, finding a complex training image representing all desired structures can be difficult in 2-D. In 3-D, the problem becomes even more acute. Nonstationarity may require using a training image of increased complexity or even multiple training images. In fact, in order to build a consistent 3-D training image, it can be necessary to perform a full geomodeling exercise, using either other stochastic simulation methods, e.g., Boolean [Boucher *et al.*, 2010; Comunian *et al.*, 2011; Maharaja, 2008] or pluri-Gaussian], or specific algorithms contained in commercial geomodeling packages. An advantage of the Boolean method is that it is possible to parameterize each facies separately, whereas our method transforms entire data events, regardless of which facies are affected. On the other hand, object-based training images have important drawbacks, such as the necessity to deal with categorical variables, and the high computational costs incurred by large and complex training images. In 2-D, it is possible to use analogues as training images, but the analog may not present the appropriate properties of stationarity and size relative to the richness of patterns. In comparison, our approach is simpler.

[57] To demonstrate the value and the practicality of our methodology, we present a 3-D synthetic aquifer modeled with elementary 3-D training images and rotation-invariant distances. Six orientation measurements are available

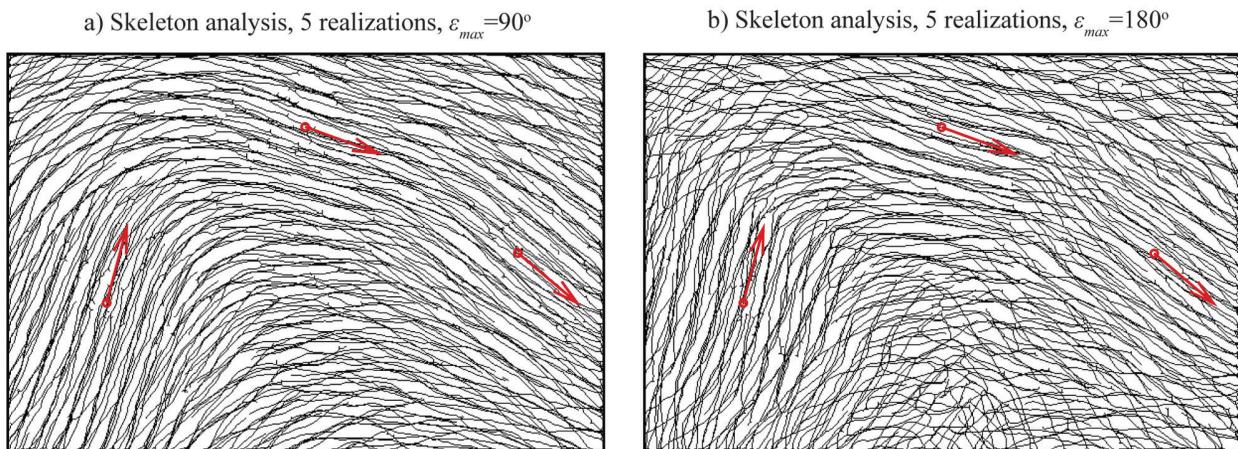


Figure 8. Skeleton analysis of realizations obtained with (a) $\varepsilon_{\max} = 90^\circ$. (b) $\varepsilon_{\max} = 180^\circ$.

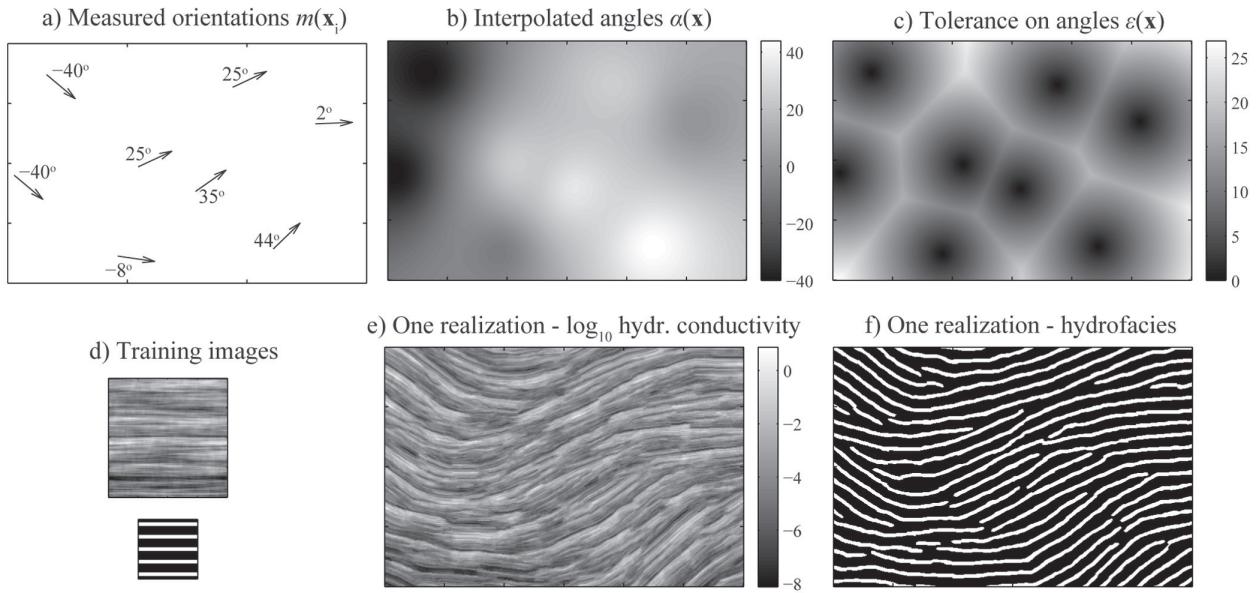


Figure 9. Application of nonstationary transformation parameters to model a double fold in 2-D. (a) Orientation measurements. (b–c) Interpolated rotation parameters. (d) Two different training images. (e–f) corresponding realizations.

(Figure 11a), which correspond to a relatively complex geological structure consisting of an anticline on one side of the domain and a syncline on the other side. As for the previous examples, we perform simulations for both categorical and continuous cases using two different training images.

[58] The categorical training image (Figure 11b) is binary and made of horizontal equally spaced layers. The continuous training image (Figure 11c) represents \log_{10} hydraulic conductivity. It has the same structure as the

categorical one, with each category populated using an unconditional multi-Gaussian realization generated by the spectral method. For both facies, exponential variograms are used with ranges of 100 units in the X and Y directions and five units in the Z direction. For the regions corresponding to facies 0 (transparent on the figures), a mean \log_{10} hydraulic conductivity of -6 is used. For facies 1 the mean \log_{10} hydraulic conductivity is set to -3. For both facies a variance of 2 was chosen. Both categorical and

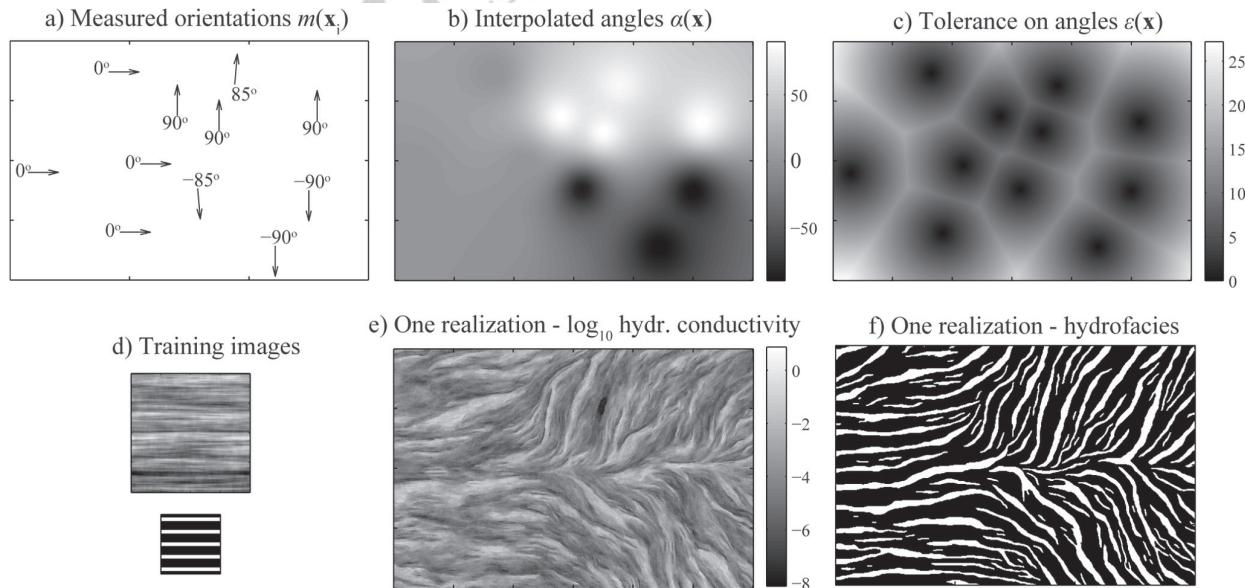


Figure 10. Application of nonstationary transformation parameters to model deltaic structures in 2-D. (a) Orientation measurements. (b–c) Interpolated rotation parameters. In addition, uniform affinity-invariant transforms are used corresponding to $A_{[1/3,3]}$. (d) Two different training images. (e–f) corresponding realizations.

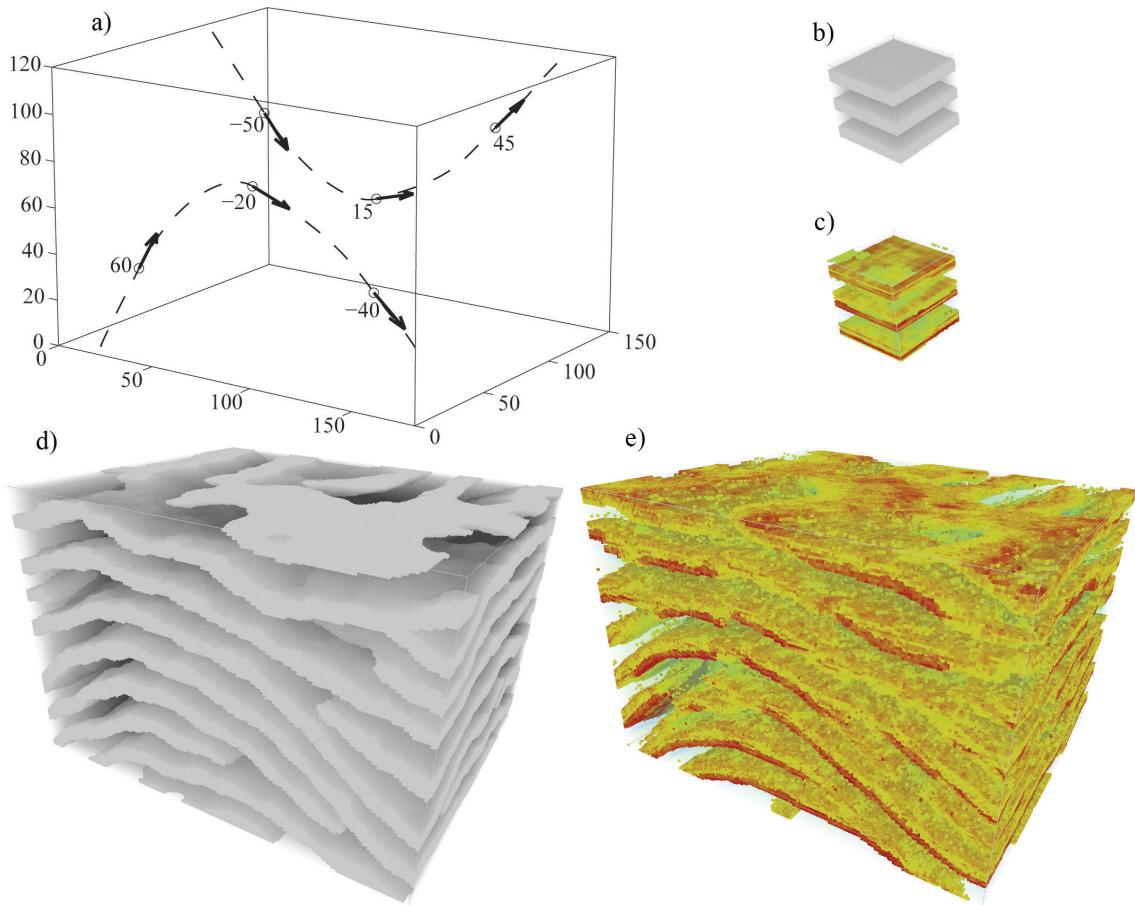


Figure 11. A high-resolution 3-D example of the use of elementary training images with transform-invariant distances. (a) Six orientation data depicting an anticline on one side of the domain and a syncline of the other side. (b–c) Elementary training images. (d–e) Corresponding realizations.

continuous training image sizes are $40 \times 40 \times 40$ grid nodes, which is small compared to the simulation domain that measures $180 \times 150 \times 120$ grid nodes.

[59] Interpolation of the angles in 3-D is performed using equation (4). The angle tolerances are defined using equation (5), with $\varepsilon_{\max} = 90^\circ$. In the direct sampling simulation, the distance threshold is set to 0.1 for the categorical simulation, 0.03 for the continuous simulation, and the number of neighbors is 25 for both cases. Figure 11d and 11e show the resulting realizations for both categorical and continuous cases. Note that the computation times are kept reasonable because of the small size of the training image. For the categorical case, the simulation of 3.24 million grid nodes takes about 1 h on a dual core 2.5 Ghz laptop, using the parallelization approach described in Mariethoz [2010] to take advantage of the computation capacity of both CPUs. Note that this result is obtained by using a relatively high distance threshold (0.1), producing simulations with satisfying large-scale structures, but that are noisy at a small scale. In a second step, the noise is removed by resimulating all pixels with a reduced neighborhood.

[60] This example demonstrates that it is suitable to use elementary training images in complex geological environments where the classical multiple-point approach would be difficult to put into practice. Our methodology is not

intended to replace structural grid transformation methods [e.g., Caers, 2005; Mallet, 2008]. However, it can be an interesting alternative for modeling structural uncertainty in cases where, although the general structures are known, one wants several alternative realizations of a structural model. This situation corresponds to the setting of the examples depicted in Figures 8 and 11.

6. Discussion and Conclusion

[61] We present a new framework for multiple-point simulation involving the use of elementary training images and transform-invariant distances. Transformations (rotations and/or affinity in the cases presented) are randomly applied to the patterns of the training image resulting in a vast family of geological structures that all display a type of geological continuity related to the training image. The transform-invariant distances are chosen with the goal of reproducing specific characteristics of the training image, such as connectivity patterns that have a strong impact on flow and transport, while leaving a degree of freedom in other characteristics such as the orientation or the anisotropy of the structures. Note that other transformations could be used, such as symmetries or shifts. The advantages of this method are that the training images are so

simple that they can be easily built, even in 3-D. After the random transformations have been applied, complex geological structures are obtained, whose spatial structures can be parameterized by adjusting the statistics of the random transformations, based on field data or geological context. Overall, our methodology is a simple way of shaping the prior spatial model, which is possible in most cases by adjusting two numbers only.

[62] We apply the method on a synthetic example involving seismic data where the transformation parameters are data-driven. We also show examples where realistic 2-D and 3-D structures are built from simplistic training images, with transformation parameters inferred using a small number of data. In the 3-D case, our method can also be used in a similar way as geomodeling tools using the implicit approach. In fact, it requires the same type of data, consisting of facies and orientation measurements.

[63] The simple parameterization of complex geological structures can be very useful in the context of inverse problems where the prior model of spatial continuity is uncertain. The transformation parameters could be inferred either with an inverse procedure or by direct adjustment to the data, as is currently done with variograms. Although these topics are not the focus of this paper, we see them as important future research areas.

[64] Another possible direction for future research would be to adapt the concept of transform-invariant distances to pattern-based simulation algorithms. The main challenge would be computational issues related to the storage requirements of the numerous transformed patterns.

[65] One fundamental issue raised by the use of elementary training images is that the training image is no longer a direct reflection of the structures deemed to exist in the subsurface, but rather a vehicle for geological representation. In fact, multiple-point simulation is often seen as a process which, given a training image, produces other images (simulations) with identical statistical properties and possibly anchored to conditioning data. However, it has been noted [Boucher, 2007; Journel and Zhang, 2006] that one cannot expect a complete statistical similarity between the training image and the corresponding simulations. In fact, the algorithmic machine lies between the training image and the simulation results. The function that produces a simulation is algorithmically defined, and hence intractable. With the same training image, different implementations of a multiple-point simulation algorithm can produce different simulation results. Each algorithm is implemented in a specific way and takes different parameters that may not have their counterpart in other implementations. In this context, a viable way of validating a training image and the associated parameters is to perform an a posteriori check to investigate whether the simulated structures correspond to what is expected by the geologist. In practice, this geological validation is largely used for both MPS and covariance-based simulation methods. More formal ways would call for inferring the training image based on field measurements, through inversion or cross-validation. This is an active and promising research topic.

[66] In this paper we acknowledge that the training image is only an initial representation of the subsurface which is further processed by the simulation algorithm. Hence, we propose an alternative paradigm where the training image

represents broad spatial concepts rather than a specific geological reality. The training images used are very simple, and complexity emerges from random transformations that are guided by conditioning data or by local orientations.

[67] **Acknowledgments.** This work was supported by the Australian Research Council, the National Water Commission, and the Swiss National Science Foundation (grant PBNEP2-124334). We thank Andre Journel, David Ginsbourger, Sanjay Srinivasan, and Juan Fernández-Martínez for fruitful brainstorming sessions and Philippe Renard, Julien Straubhaar, Jef Caers, and two anonymous reviewers for their constructive comments.

References

- Allard, D., R. Froidevaux, and P. Biver (2006), Conditional simulation of multi-type non stationary Markov object models respecting specified proportions, *Math. Geosci.*, 38(8), 959–986, doi:10.1007/s11004-006-9057-5.
- Arpat, B., and J. Caers (2007), Conditional simulations with patterns, *Math. Geol.*, 39(2), 177–203.
- Bárdossy, A., and J. Li (2008), Geostatistical interpolation using copulas, *Water Resour. Res.*, 44, W07412, doi:10.1029/2007WR006115.
- Bastante, F., C. Ordóñez, J. Taboada, and J. Matías (2008), Comparison of indicator kriging, conditional indicator simulation and multiple-point statistics used to model slate deposits, *Eng. Geol.*, 98(1–2), 50–59, doi:10.1016/j.enggeo.2008.01.006.
- Boucher, A. (2007), Algorithm-driven and representation-driven random function: A new formalism for applied geostatistics, edited, Stanford Center for Reservoir Forecasting, Stanford, CA.
- Boucher, A., R. Gupta, J. Caers, and A. Satija (2010), Tetris: A training image generator for SGeMS, paper presented at Proceedings of 23rd SCRF annual affiliates meeting, Stanford University, May 5–6 2010.
- Caers, J. (2003), History matching under a training image-based geological model constraint, *SPE J.*, 8(3), 218–226, SPE # 74716.
- Caers, J. (2005), *Petroleum Geostatistics*, 88 pp., Society of Petroleum Engineers, Richardson, TX.
- Chugunova, T., and L. Hu (2006), Multiple-Point Statistical Simulations Constrained by continuous Auxiliary Data, paper presented at Int. Assoc. for Math. Geol. Xlth International Congress, Université de Liège, Belgium.
- Cojan, I., O. Fouche, S. Lopez, and J. Rivoirard (2004), Process-based reservoir modelling in the example of meandering channel, in *Geostatistics Banff*, edited by O. Leuangthong and C. Deutch, pp. 611–619, Springer, Berlin, Germany.
- Comunian, A., P. Renard, J. Straubhaar, and P. Bayer (2011), Three-dimensional high resolution fluvio-glacial aquifer analog: Part 2: geostatistical modeling, *J. Hydrol.*, doi:10.1016/j.jhydrol.2011.03.037, in press.
- de Almeida, J. (2010), Stochastic simulation methods for characterization of lithoclasses in carbonate reservoirs, *Earth Sci. Rev.*, 101(3–4), 250–270.
- De Vries, L., J. Carrera, O. Falivene, O. Gratacos, and L. Slooten (2009), Application of multiple point geostatistics to non-stationary images, *Math. Geosci.*, 41(1), 29–42.
- Deutsch, C., and L. Wang (1996), Hierarchical object-based stochastic modeling of fluvial reservoirs, *Math. Geol.*, 28(7), 857–880, doi:10.1007/BF02066005.
- Deutsch, C., and T. Tran (2002), FLUVSIM: A program for object-based stochastic modeling of fluvial depositional systems, *Comput. Geosci.*, 28(4), 525–535.
- Dimitrakopoulos, R., H. Mustapha, and E. Gloaguen (2010), High-order statistics of spatial random fields: Exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena, *Math. Geosci.*, 42(1), 65–99.
- Efros, A., and T. Leung (1999), Texture synthesis by non-parametric sampling, paper presented at The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, Kerkyra, Greece.
- Emery, X. (2007), Using the Gibbs sampler for conditional simulation of Gaussian-based random fields, *Comput. Geosci.*, 33(2007), 522–537.
- Emery, X., and J. Ortiz (2011), A comparison of random field models beyond bivariate distributions, *Math. Geosci.*, 43(2), 183–202, doi:10.1007/s11004-010-9305-6.
- Feyen, L., and J. Caers (2006), Quantifying geological uncertainty for flow and transport modelling in multi-modal heterogeneous formations, *Adv. Water Resour.*, 29(6), 912–929.

- 905 Gómez-Hernández, J. J., and X.-H. Wen (1998), To be or not to be multi- 971
gaussian? A reflection on stochastic hydrogeology, *Adv. Water Resour.*, 972
21(1), 47–61.
906 Gonzales, E., T. Mukerji, and G. Mavko (2008), Seismic inversion combining 973
rock physics and multiple-point geostatistics, *Geophysics*, 73(1), 974
R11–R21, doi:10.1190/1.2803748.
907 Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, 975
Oxford University Press, Oxford, U.K.
908 Green, C., J. Böhlke, B. Bekins, and S. Philips (2010), Mixing effects on 976
apparent reaction rates and isotope fractionation during denitrification in 977
a heterogeneous aquifer, *Water Resour. Res.*, 46, W08525, doi:10.1029/ 978
2009WR008903.
909 Gumiaux, C., D. Gapais, and J. Brun (2003), Geostatistics applied to 979
best-fit interpolation of orientation data, *Tectonophysics*, 376(3–4), 980
241–259.
910 Honarkhah, M., and J. Caers (2010), Stochastic simulation of patterns 981
using distance-based pattern modeling, *Math. Geosci.*, 42(5), 487–517, 982
doi:10.1007/s11004-010-9276-7.
911 Hu, L., and T. Chugunova (2008), Multiple-point geostatistics for modeling 983
subsurface heterogeneity: A comprehensive review, *Water Resour. Res.*, 984
44, W11413, doi:10.1029/2008WR006993.
912 Huysmans, M., and A. Dassargues (2009), Application of multiple-point 985
geostatistics on modelling groundwater flow and transport in a cross- 986
bedded aquifer (Belgium), *Hydrogeol. J.*, 17, 1901–1911.
913 Journel, A., and T. Zhang (2006), The necessity of a multiple-point prior 987
model, *Math. Geol.*, 38(5), 591–610.
914 Keogh, K., A. Martinius, and R. Osland (2007), The development of fluvial 988
stochastic modelling in the Norwegian oil industry: A historical review, 989
subsurface implementation and future directions, *Sediment. Geol.*, 990
202(1–2), 249–268.
915 Kitanidis, P. (1995), Quasi-linear geostatistical theory for inverting, *Water 991
Resour. Res.*, 31(10), 2411–2419, doi:10.1029/95WR01945.
916 Klise, K., G. Weissmann, S. McKenna, E. Nichols, J. Frechette, T. Wawrzyniec, 992
and V. Tidell (2009), Exploring solute transport and streamline 993
connectivity using lidar-based outcrop images and geostatistical representations 994
of heterogeneity, *Water Resour. Res.*, 45, W05413, doi:10.1029/ 995
2008WR007500.
917 Knudby, C., and J. Carrera (2005), On the relationship between indicators 996
of geostatistical, flow and transport connectivity, *Adv. Water Resour.*, 997
28(4), 405–421.
918 Lajaunie, C., G. Courrioux, and L. Manuel (1997), Foliations field and 3D 998
cartography in geology: Principles of a method based on potential interpolation, 999
Math. Geol., 29(4), 571–584, doi:10.1007/BF02775087.
920 Lantuéjoul, C. (2002), *Geostatistical simulation: Models and Algorithms*, 1000
232 pp., Springer, Berlin, Germany.
921 Le Loc'h, G., and A. G. Galli (1994), Improvement in the truncated Gaussian 1001
method: Combining several Gaussian functions, paper presented at 1002
Ecmor 4, 4th European Conference on the Mathematics of Oil Recovery, 1003
Roros, Norway, 7–10 June.
922 Le Ravalec-Dupin, M., B. Noetinger, and L. Y. Hu. (2000), The FFT moving 1004
average (FFT-MA) generator: An efficient numerical method for 1005
generating and conditioning Gaussian simulations, *Math. Geol.*, 32(6), 1006
701–723.
923 Lu, D., T. Zhang, J. Yamg, D. Li, and X. Kong (2009), A reconstruction 1007
method of porous media integrating soft data with hard data, *Chin. Sci. Bull.*, 1008
54(11), 1876–1885.
924 Maharaja, A. (2008), TiGenerator: Object-based training image generator, 1009
Comput. Geosci., 34(12).
925 Mallet, J.-L. (2008), *Numerical Earth Models*, EAGE Pub. BV, Houten, 1010
Netherlands.
926 Mariethoz, G., P. Renard, F. Cornaton, and O. Jaquet (2009), Truncated 1011
plurigaussian simulations to characterize aquifer heterogeneity, *Ground 1012
Water*, 47(1), 13–24, doi:10.1111/j.1745-6584.2008.00489.x.
927 Mariethoz, G. (2010), A general parallelization strategy for random path 1013
based geostatistical simulation methods, *Comput. Geosci.*, 37(7), 953– 1014
958, doi:10.1016/j.cageo.2009.11.001.
928 Mariethoz, G., and P. Renard (2010), Reconstruction of incomplete data 1015
sets or images using Direct Sampling, *Math. Geosci.*, 42(3), 245–268, 1016
doi:10.1007/s11004-010-9270-0.
929 Mariethoz, G., P. Renard, and J. Straubhaar (2010), The direct sampling 1017
method to perform multiple-point simulations, *Water Resour. Res.*, 46, 1018
W11536, doi:10.1029/2008WR007621.
930 Maxelon, M., and N. Mancktelow (2005), Three-dimensional geometry 1019
and tectonostratigraphy of the Pennine zone, Central Alps, Switzerland 1020
and Northern Italy, *Earth Sci. Rev.*, 71(3–4), 171–227, doi:10.1016/ 1021
j.earscirev.2005.01.003.
931 Maxelon, M., P. Renard, G. Courrioux, M. Brändli, and N. Mancktelow 1022
(2009), A workflow to facilitate three-dimensional geometrical modeling 1023
of complex poly-deformed geological units, *Comput. Geosci.*, 35(3), 1024
644–658, doi:10.1016/j.cageo.2008.06.005.
932 Michael, H., A. Boucher, T. Sun, J. Caers, and S. Gorelick (2010), Combin- 1025
ing geologic-process models and geostatistics for conditional simulation 1026
of 3-D subsurface heterogeneity, *Water Resour. Res.*, 46, W05527, 1027
doi:10.1029/2009WR008414.
933 Nowak, W., F. de Barros, and Y. Rubin (2010), Bayesian geostatistical 1028
design: Task-driven optimal site investigation when the geostatistical 1029
model is uncertain, *Water Resour. Res.*, 46, W03535, doi:10.1029/ 1030
2009WR008312.
934 Okabe, H., and M. Blunt (2007), Pore space reconstruction of vuggy carbonates 1031
using microtomography and multiple-point statistics, *Water 1032
Resour. Res.*, 43, W12S02, doi:10.1029/2006WR005680.
935 Pyrcz, M., J. Boisvert, and C. V. Deutsch (2008), A library of training 1033
images for fluvial and deepwater reservoirs and associated code, *Comput. 1034
Geosci.*, 34(5), 542–560.
936 Pyrcz, M., J. Boisvert, and C. Deutsch (2009), ALLUVSIM: A program for 1035
event-based stochastic modeling of fluvial depositional systems, *Comput. 1036
Geosci.*, 35(8), 1671–1685, doi:10.1016/j.cageo.2008.09.012.
937 Ronayne, M., S. Gorelick, and J. Caers (2008), Identifying discrete geologic 1037
structures that produce anomalous hydraulic response: An inverse modeling 1038
approach, *Water Resour. Res.*, 44, W08426, doi:10.1029/2007WR006635.
938 Sánchez-Vila, X., J. Carrera, and J. P. Girardi (1996), Scale effects in trans- 1039
missivity, *J. Hydrol.*, 183(1–2), 1–22.
939 Straubhaar, J., P. Renard, G. Mariethoz, R. Froidevaux, and O. Besson 1040
(2011), An improved parallel multiple-point algorithm using a list 1041
approach, *Math. Geosci.*, 43(3), 305–328, doi:10.1007/s11004-011-9328-7.
940 Strebelle, S. (2002), Conditional simulation of complex geological struc- 1042
tures using multiple-point statistics, *Math. Geol.*, 34(1), 1–22.
941 Strebelle, S., and T. Zhang (2004), Non-stationary multiple-point geo- 1043
statistical models, in *Geostatistics Banff*, edited by O. Leuangthong 1044
and C. Deutch, pp. 235–244, Springer, Berlin, Germany.
942 Suzuki, S., and J. Caers (2008), A distance-based prior model parameteriza- 1045
tion for constraining solutions of spatial inverse problems, *Math. Geosci.*, 1046
40(4), 445–469.
943 Western, A., G. Blöschl, and R. Grayson (2001), Toward capturing hydro- 1047
logically significant connectivity in spatial patterns, *Water Resour. Res.*, 1048
37(1), 83–97, doi:10.1029/2000WR900241.
944 Wojciek, R., D. McLaughlin, A. Konings, and D. Entekhabi (2009), Condi- 1049
tioning stochastic rainfall replicates on remote sensing data, *IEEE Trans. 1050
Geosci. Remote Sens.*, 47(8), 2436–2449.
945 Wu, J., A. Boucher, and T. Zhang (2008), A SGEMS code for pattern simu- 1051
lation of continuous and categorical variables: FILTERSIM, *Comput. 1052
Geosci.*, 34(12), 1863–1876.
946 Zhang, T., P. Switzer, and A. Journel (2006), Filter-based classification of 1053
training image patterns for spatial simulation, *Math. Geol.*, 38(1), 63–80.
947 Zinn, B., and C. F. Harvey (2003), When good statistical models of aquifer 1054
heterogeneity go bad: A comparison of flow, dispersion, and mass trans- 1055
fer in connected and multivariate Gaussian hydraulic conductivity fields, 1056
Water Resour. Res., 39(3), 1051, doi:10.1029/2001WR001146.
948 B. F. J. Kelly and G. Mariethoz, National Centre for Groundwater Research 1057
and Training, University of New South Wales, Sydney, Australia. (gregoire.mariethoz@minds.ch)
949 1058
950 1059
951 1060
952 1061
953 1062
954 1063
955 1064
956 1065
957 1066
958 1067
959 1068
960 1069
961 1070
962 1071
963 1072
964 1073
965 1074
966 1075
967 1076
968 1077
969 1078
970 1079