Mathematical Geology, Vol. 25, No. 3, 1993

Entropy and Spatial Disorder[1]

André G. Journel[2] and Clayton V. Deutsch[2]

The majority of geostatistical estimation and simulation algorithms rely on a covariance model as the sole characteristic of the spatial distribution of the attribute under study. The limitation to a single covariance implicitly calls for a multivariate Gaussian model for either the attribute itself or for its normal scores transform. The Gaussian model could be justified on the basis that it is both analytically simple and it is a maximum entropy model, i.e., a model that minimizes unwarranted structural properties. As a consequence, the Gaussian model also maximizes spatial disorder (beyond the imposed covariance) which can cause flow simulation results performed on multiple stochastic images to be very similar; thus, the space of response uncertainty could be too narrow entailing a misleading sense of safety. The ability of the sole covariance to adequately describe spatial distributions for flow studies, and the assumption that maximum spatial disorder amounts to either no additional information or a safe prior hypothesis are questioned. This paper attempts to clarify the link between entropy and spatial disorder and to provide, through a detailed case study, an appreciation for the impact of entropy of prior random function models on the resulting response distributions.

KEY WORDS: spatial entropy, Gaussian model, simulation, fluid flow.

INTRODUCTION

Although extensively used in thermodynamics and information theory (Kullback, 1968; Jaynes, 1985a,b), the concept of entropy has only been recently introduced in geostatistics (Christakos, 1990). Christakos (1990) introduces entropy as a measure of the uncertainty of a prior distribution model; the principle is to maximize that uncertainty (entropy) beyond the statistics that are considered known. The intent is that the prior random function model should only account for known statistics. If those statistics reduce to the covariance function then, it can be shown that the maximum entropy random function model is Gaussian. However, whenever enough data are available to allow inference of a covariance model they also usually provide valuable additional information; in which case the maximum entropy prior random function model may no longer be Gaussian.

A case study will demonstrate that alternative prior random function models

[1]Received 12 May 1992; accepted 31 July 1992.

[2]Department of Applied Earth Sciences, Stanford University, Stanford, California 94305.

0882-8121/93/0400-0329$07.00/1

sharing the same covariance model could lead to much larger uncertainty in the response variables of transport simulations, such as effective permeability, breakthrough times, or oil/water sweepage.

ENTROPY OF CONTINUOUS DISTRIBUTIONS

Let $f_z = f_z(z_1, z_2, \ldots, z_n)$ be the n-variate probability density function (pdf) of the n random variables (RVs) Z_i, $i = 1, \ldots, n$. The uncertainty associated to the distribution f_z is measured by its entropy defined as (Shannon, 1948; Christakos, 1990):

$$H_{f_z} = E\{-\ln f_z(z_1, \ldots, z_n)\}$$

$$= -\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} [\ln f_z(z_1, \ldots, z_n)] f_z(z_1, \ldots, z_n)\, dz_1 \ldots dz_n \quad (1)$$

Since the function $-\ln(y)$ increases as y decreases, the smaller the pdf value $f_z(z_1, \ldots, z_n)$ the larger its contribution to the entropy measure H. The greater the uncertainty, the more spread and the smaller the pdf values f_z, the larger the entropy.

Univariate pdf

For $n = 1$, Eq. (1) defines the entropy of a univariate continuous distribution with pdf $f(z)$:

$$H_f = -\int_{-\infty}^{+\infty} [\ln f(z)] f(z)\, dz \quad (2)$$

All intervals within which the pdf $f(z)$ is zero are excluded from the integral. Classical results include (Shannon, 1948; Jones, 1979):

- The bounded pdf with maximum entropy is the uniform distribution, i.e., $f(z) = 1/(b - a)$ for $z \in [a, b]$ and $f(z) = 0$ otherwise. The entropy is $H_f = \ln(b - a)$ which decreases to $-\infty$ as the interval $[a, b]$ becomes narrower corresponding to greater certainty.
- For a fixed variance, the unbounded pdf that maximizes entropy is the normal (Gaussian) distribution. In practice, most histograms are bounded and very few can be considered as normal. However, a normal score transform can be applied and the previous result holds for the normal pdf of the transforms.

Bivariate pdf

For $n = 2$, or for two random variables $Z(\mathbf{u})$ and $Z(\mathbf{u} + \mathbf{h})$ separated by a given vector $\mathbf{h}$, the entropy associated to the stationary bivariate pdf $f_{\mathbf{h}}(z, z')$ is:

$$H_f(\mathbf{h}) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [\ln f_{\mathbf{h}}(z, z')] f_{\mathbf{h}}(z, z')\, dz\, dz' \tag{3}$$

The previous univariate results extend to the bivariate case:

- For bounded random variables, maximum entropy is obtained for the uniform or rectangular bivariate pdf.
- Among all bivariate pdfs sharing the same covariance function $C(\mathbf{h}) = \text{Cov}\{Z(\mathbf{u}), Z(\mathbf{u} + \mathbf{h})\}$ or the same covariance matrix $E\{(\mathbf{z} - \mathbf{m})(\mathbf{z} - \mathbf{m})^T\}$, the Gaussian pdf maximizes entropy. Thus, if the only prior structural information (spatial statistics) retained is the covariance function $C(\mathbf{h})$, the maximum entropy random function (RF) model is the Gaussian model.
- Different bivariate pdfs may share the same covariance $C(\mathbf{h})$ and yet have different entropy functions $H_f(\mathbf{h})$, or they may share both covariance and entropy functions. One could also imagine different bivariate pdfs with the same bivariate entropy and different covariance functions.

ENTROPY OF DISCRETE DISTRIBUTIONS

The integral Eqs. (2) and (3) of entropy call for continuous pdfs and are used only for analytical developments. In most practical applications, the variable $Z(\mathbf{u})$ is either categorical (e.g., rock types) or, if continuous, is discretized into a finite number of classes, e.g., the classes of its histogram.

Let Z be a discrete RV that can take K outcome values (or be valued in K predefined classes) with probabilities p_k, $k = 1, \ldots, K$, such that $\sum_{k=1}^{K} p_k = 1$.

The entropy attached to this discrete probability set is defined as:

$$H = -\sum_{k=1}^{K} [\ln p_k] p_k \geq 0 \tag{4}$$

Similarly, consider the stationary discrete RF $Z(\mathbf{u})$ with for stationary marginal probabilities the set $\{p_k, k = 1, \ldots, K\}$, and for bivariate probabilities the set:

$$p_{k,k'}(\mathbf{h}) = \text{Prob}\{Z(\mathbf{u}) \in \text{category } k,\ Z(\mathbf{u} + \mathbf{h}) \in \text{category } k'\}$$

$$\text{independent of } \mathbf{u};\ k, k' = 1, \ldots, K \tag{5}$$

Note that $\sum_{k'}^{K} p_{k,k'}(\mathbf{h}) = p_k = \text{Prob}\{Z(\mathbf{u}) \in \text{category } k\}$, for all $\mathbf{h}$.

The entropy associated to that set of bivariate probabilities is defined as:

$$H(\mathbf{h}) = -\sum_{k=1}^{K}\sum_{k'=1}^{K} [\ln p_{k,k'}(\mathbf{h})] p_{k,k'}(\mathbf{h}) \geq 0 \tag{6}$$

Some remarks:

- The discrete case entropy is nonnegative because discrete probabilities can not be greater than 1, whereas in the continuous case the entropy can be as low as $-\infty$.
- If the RV $Z(\mathbf{u})$ is not naturally categorical but is made discrete through classification, that class definition impacts the entropy value as defined in Eqs. (4) or (6). Hence, before comparing the discrete entropies of originally continuous distributions, care should be taken to standardize the class definition.
- For lag $\mathbf{h} = 0$: $p_{k,k'}(0) = 0$, $\forall k \neq k'$, and: $p_{kk}(0) = p_k$. Thus, Eq. (6) entails:

$$H(0) = -\sum_{k=1}^{K} [\ln p_k] p_k \equiv H \tag{7}$$

The univariate entropy represents the lower bound of the bivariate entropy, i.e., $H(\mathbf{h}) \geq H(0)$, $\forall \mathbf{h}$. This lower bound can be seen as the case corresponding to perfect correlation: $Z(\mathbf{u} + \mathbf{h}) \equiv Z(\mathbf{u})$.

- For lag $\mathbf{h} = +\infty$, and two very distant RVs $Z(\mathbf{u})$ and $Z(\mathbf{u} + \mathbf{h})$ can be considered as independent, hence:

$$p_{k,k'}(+\infty) = p_k p_{k'}, \ \forall k, k'$$

and from Eq. (6):

$$H(\infty) = -\sum_{k=1}^{K} \sum_{k'=1}^{K} [\ln p_k + \ln p_{k'}] p_k p_{k'}$$

$$= -\sum_{k=1}^{K} p_k \ln p_k \sum_{k'=1}^{K} p_{k'} - \sum_{k'=1}^{K} p_{k'} \ln p_{k'} \sum_{k=1}^{K} p_k = 2H(0)$$

$$H(\infty) = 2H(0) \tag{8}$$

For most well-behaved nonperiodic pdfs, for which $p_{k,k'}(\mathbf{h})$ decreases continuously as $|\mathbf{h}|$ increases along any direction, the bivariate entropy measure $H(\mathbf{h})$ will increase with $|\mathbf{h}|$. For this class of bivariate cdfs, the entropy measure $H(\mathbf{h})$ is increasing in the interval $[H(0), H(+\infty)]$, i.e.,

$$H(\mathbf{h}) \in [H(0), 2H(0)] \tag{9}$$

Thus, a standardized relative measure of bivariate entropy is:

$$H_R(\mathbf{h}) = \frac{H(\mathbf{h}) - H(0)}{H(0)} \in [0, 1] \tag{10}$$

The upper bound (maximum entropy) $H_R(\infty) = 1$ corresponds to the case of two independent variables, i.e., to maximum unpredictability, which results

from maximum spatial "disorder" or "disorganization." On the other hand, if a particular class of z-values, say, k_0 corresponding to small z-values, persists over long distances $\mathbf{h}$, then the distribution of z in space presents a feature or "spatial order" that helps predicting unsampled z-values. As defined above, spatial disorder increases with the entropy (uncertainty) measure $H(\mathbf{h})$ or $H_R(\mathbf{h})$.

The measure $H_R(\mathbf{h})$ is to be compared with the relative semi-variogram, a measure of (bivariate) spatial variability defined as:

$$\gamma_R(\mathbf{h}) = \frac{1}{2}\frac{E\{[Z(\mathbf{u}+\mathbf{h}) - Z(\mathbf{u})]^2\}}{C(0)} = \frac{C(0) - C(\mathbf{h})}{C(0)} \in [0, 1] \qquad (11)$$

with $C(\mathbf{h})$ being the stationary covariance function Cov $\{Z(\mathbf{u}), Z(\mathbf{u}+\mathbf{h})\}$ and $C(0)$ the corresponding variance.

SPATIAL ENTROPY

Building on the similarity of Eqs. (10) and (11), one can define an average bivariate entropy over a field A of measure $|A|$ as:

$$\overline{H_R}(A, A) = \frac{1}{|A|^2}\int_A d\mathbf{u}\int_A H_R(\mathbf{u} - \mathbf{u}')\, d\mathbf{u}' \in [0, 1] \qquad (12)$$

In practice the field A would be discretized by n locations of coordinates $\mathbf{u}_i$, $i = 1, \ldots, n$, and the space integral (12) approximated by:

$$\overline{H_R}(A, A) \cong \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} H_R(\mathbf{u}_i - \mathbf{u}_j) \qquad (13)$$

When $A \to \infty$, $\overline{H_R}(A, A) \to 1$, since $H_R(\infty) \to 1$.

Equations (12) and (13) are similar to the expressions defining the dispersion (expected spatial) variance of $Z(\mathbf{u})$ within the field A, (Journel and Huijbregts, 1978, p. 67):

$$D^2(0/A) = E\{S^2(0/A)\} = \bar{\gamma}(A, A) = \frac{1}{|A|^2}\int_A d\mathbf{u}\int_A \gamma(\mathbf{u} - \mathbf{u}')\, d\mathbf{u}' \qquad (14)$$

with $S^2(0/A)$ being the randomization of the spatial variance

$$s^2(0/A) = \frac{1}{|A|}\int_A [z(\mathbf{u}) - m_A]^2\, d\mathbf{u}, \quad \text{and} \quad m_A = \frac{1}{|A|}\int_A z(\mathbf{u})\, d\mathbf{u}$$

The average relative entropy measure $\overline{H_R}(A, A)$, defined in (12) could be used as a global measure of bivariate spatial entropy (disorder) over the field A. However, whether this measure will prove more useful than the little used dispersion variance (14) is yet questionable.

Note that the measure $\overline{H_R}(A, A)$ utilizes only bivariate entropy when a full

measure of spatial entropy should use a multivariate entropy function. However, just as multivariate pdfs are rarely accessible in practice, multivariate entropy functions are not likely to be practical.

CASE STUDY

In the following case study, an *exhaustively* sampled reference image will be reproduced through stochastic simulations based on RF models differing by the statistics which constrain them. It will be shown that different models though sharing the same covariance function $C(\mathbf{h})$ and the same univariate pdf lead to realizations with different spatial entropy (spatial disorder) and, more importantly, result in a different assessment of the response uncertainty.

Figure 1 is an image of cross-stratified sands and silty sands from a distributary-mouth bar taken from p. 151 of *Sandstone Depositional Environments* (Scholle and Spearing, 1982). The grayness of the image has been arbitrarily scaled through a monotonic function into permeability values to yield the discrete histogram of Fig. 2. With the discretization shown with 1.0 millidarcy classes, the entropy of that histogram is, according to Eqs. (4) or (7):

$$H(0) = 2.035$$

For the rest of this study, the data and image of Fig. 1 will be considered as the "true" reference to be reproduced.

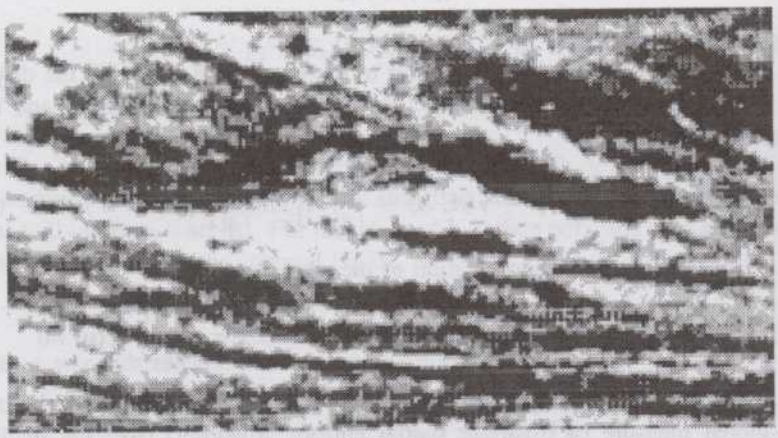

Fig. 1. Original core image: a scanned photograph of core from a distributary-mouth bar sequence (from Scholle and Spearing, 1982, p. 151). The darker gray levels correspond to a greater proportion of clayey/silty material. The actual image is approximately 6 inches by 3 inches.

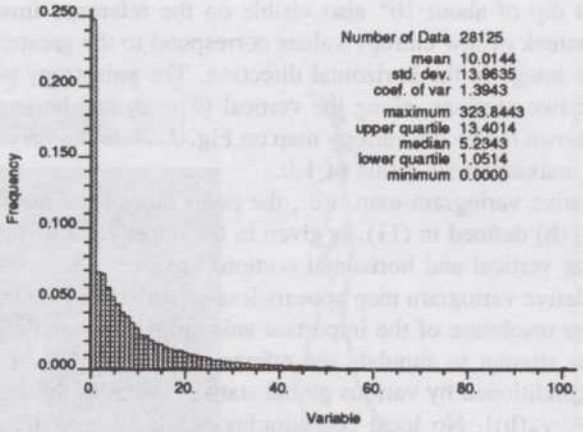

Fig. 2. Distribution of the reference permeability values (md).

The reference image of Fig. 1 is discretized into 225 × 125 square pixels. For each separation vector **h**, all pairs of pixels separated by **h** have been scanned and the exhaustive proportions $p_{k,k'}(\mathbf{h})$ defined in (5) have been calculated; the categories k, k' correspond to classes of 1 md interval. The corresponding experimental (*exhaustive*) bivariate entropy $H(\mathbf{h})$, as defined in (6), has been calculated. For example, for a unit vertical (azimuth = 0) lag, the bivariate entropy is:

$$H(1, \theta = 0) = 3.247$$

The corresponding relative entropy, from Eq. (10), is:

$$H_R(1, \theta = 0) = 0.60$$

This represents a substantial increase from the minimum value 0 in the interval [0, 1].

The exhaustive set of relative entropy values $H_R(\mathbf{h})$ for all vectors **h** within a square area A with side 100 lag units (pixels) is plotted in the upper left of Fig. 3. The plot gives $H_R(\mathbf{h}) = H_R(|\mathbf{h}|, \theta)$ in polar coordinates, with $|\mathbf{h}|$ being the modulus and θ the azimuth of **h;** the origin $H_R(\mathbf{0}) = 0$ is plotted at the center of the figure.

The spatial average (13) of the relative entropy over the square area A is (see also Table I):

$$\overline{H_R}(A, A) = 0.953$$

As discussed above this value represents a summary of bivariate spatial entropy (spatial disorder) over a window of size A.

Note the strong vertical to horizontal anisotropy shown on the entropy map

with a slight dip of about 10° also visible on the reference image of Fig. 1. The central streak of low entropy values correspond to the greater continuity of the reference image in the horizontal direction. The anisotropy is further illustrated on the two sections, along the vertical ($\theta = 0$) and horizontal ($\theta = 90$) directions, shown below the entropy map on Fig. 3. Note the very steep increase to the stable maximum sill value of 1.0.

The relative variogram map, i.e., the polar plot of the measure of spatial variability $\gamma_R(\mathbf{h})$ defined in (11), is given in the upper right of Fig. 3. The two corresponding vertical and horizontal sections are shown below the variogram map. The relative variogram map appears less smooth than the entropy map and offers a better resolution of the important anisotropy seen on Fig. 1.

We now attempt to simulate the reference image of Fig. 1 using various RF models conditioned by various global statistics such as the covariance $C(\mathbf{h}) = C(0)[1 - \gamma_R(\mathbf{h})]$. No local conditioning has been considered, that is, no samples are retained from the reference image.

In an actual study, the simulations must be made conditional to local data, whether hard (e.g., wells) or soft (e.g., seismic). The point of this paper, however, is the comparison of results obtained from RF models with different spatial entropy. This point is better made if not blurred by the effects of conditioning. Indeed, if conditioning is dense enough all realizations of any RF model would look alike, regardless of the entropy or any other characteristic of the RF model.

Gaussian Model

The reference permeability values are first transformed into a standard normal distribution through a normal scores transform. Then a multivariate Gaussian RF model is adopted for these normal scores transforms. This Gaussian RF model is fully characterized by the variogram of the normal scores transforms shown in Fig. 4. The experimental (exhaustive) variogram of the reference normal scores has been modeled by the nested sum of three structures, a nugget effect of 0.20, a short scale exponential structure contributing 0.55 of the total unit variance and a larger scale spherical structure contributing 0.25; both structures are anisotropic:

$$\gamma(h_x, h_y) = 0.20 + 0.55\ \mathrm{Exp}\left(\sqrt{\left(\frac{h_x}{10}\right)^2 + \left(\frac{h_y}{1.6}\right)^2}\right) + 0.25\ \mathrm{Sph}\left(\sqrt{\left(\frac{h_x}{180}\right)^2 + \left(\frac{h_y}{15}\right)^2}\right) \tag{15}$$

with $\mathrm{Exp}(h) = 1 - \exp(-h)$ and $\mathrm{Sph}(h) = 1.5h - 0.5h^3$, for $h \leq 1$; $= 1$ otherwise. h_x and h_y are the horizontal and vertical coordinates.

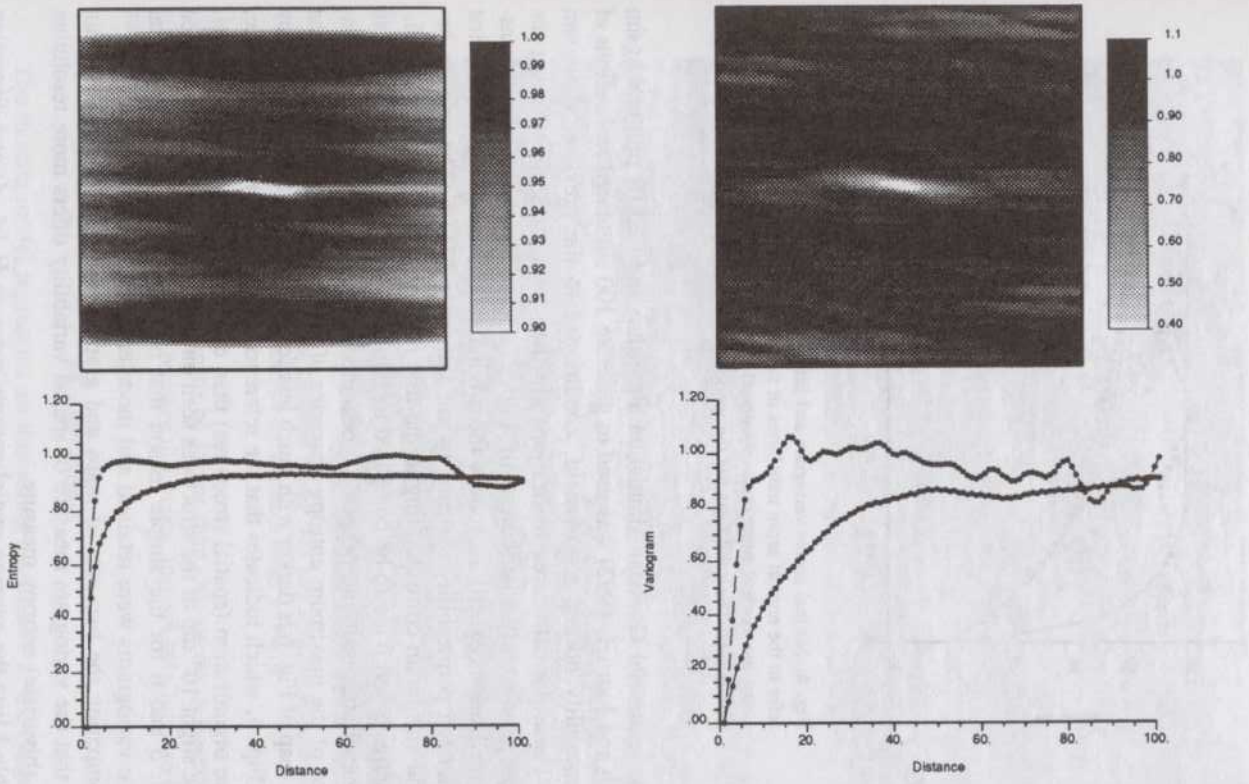

Fig. 3. Relative entropy and variogram of the reference image. The upper left map is that of bivariate entropy, the lower left graph shows the vertical and horizontal cross-sections through the entropy map. The upper right map is that of the relative variogram, and the lower right graph shows cross-sections through the variogram map.

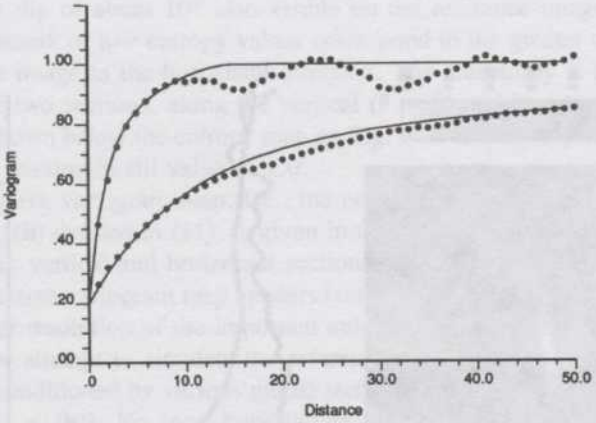

Fig. 4. Normal scores variogram and fitted model. The variograms refer to the normal score transforms of the reference data of Fig. 1 along the vertical (upper set of curves) and horizontal directions (lower set). The solid lines are the model fits.

The sequential Gaussian simulation algorithm and GSLIB program sgsim (Deutsch and Journel, 1992) was used to generate 100 simulated realizations of the permeability normal scores field, conditioned to the previous variogram model. These normal scores realizations are then back-transformed using the reference permeability distribution of Fig. 2. Consequently, all back-transformed realizations exactly reproduce the original histogram of Fig. 2. The first two generated permeability realizations are shown on Fig. 5. The relative entropy and variogram corresponding to the first realization are given on Fig. 6.

Figures 5 and 6 are to be compared to Figures 1 and 3, respectively. Note the characteristic "salt-and-pepper" appearance of the Gaussian images, a consequence of the maximum entropy character of the underlying RF model. The entropy map of Fig. 6 is darker with much less features that the reference entropy map of Fig. 3, which indicates that the reference image, beyond its covariance, has more organization (spatial structures) than exposed by the Gaussian model.

The slight 10° dip of the structures seen on Figs. 1 and 3 is not reproduced on Figs. 5 and 6 for the simple reason that only the vertical and horizontal reference variograms were retained and modeled.

Comparing the two sets of maps and graphs on Fig. 6, it is once again evident that the variogram measure of spatial variability offers more resolution that the (bivariate) entropy measure.

Table I lists the average spatial entropy measure $\overline{H_R}(A, A)$ and dispersion variance $\overline{\gamma_R}(A, A)$, for the reference image and the Gaussian realizations. "A" represents a square area of side 100 unit lags. The average entropy, $\overline{H_R}(A, A)$,

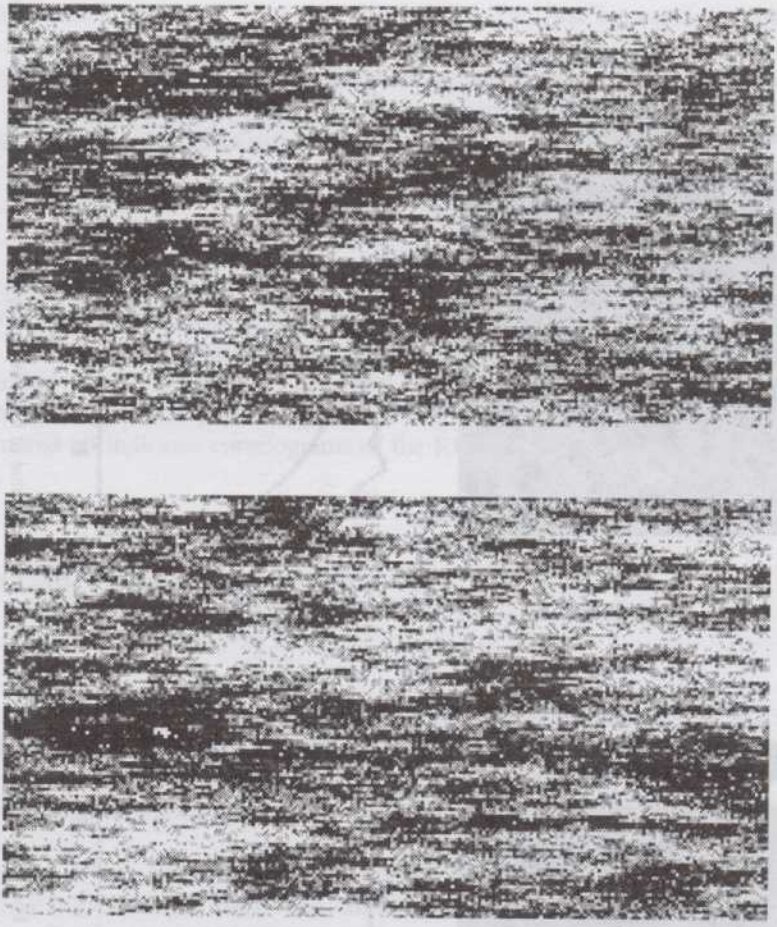

Fig. 5. Two realizations from the Gaussian model. The unit scale, in millidarcies, is the same as that used for the reference image of Fig. 1.

of the Gaussian image 1 is seen to be larger than that of the reference image ($0.994 > 0.953$), although the relative difference appears small due to the poor resolution of the entropy measure.

Mosaic Model

The mosaic model represents an interesting alternative to the Gaussian RF model in that it is also fully characterized by a single covariance function, yet its multivariate distribution is non-Gaussian.

The "mosaic" model is a bivariate mixture of two binormal distributions

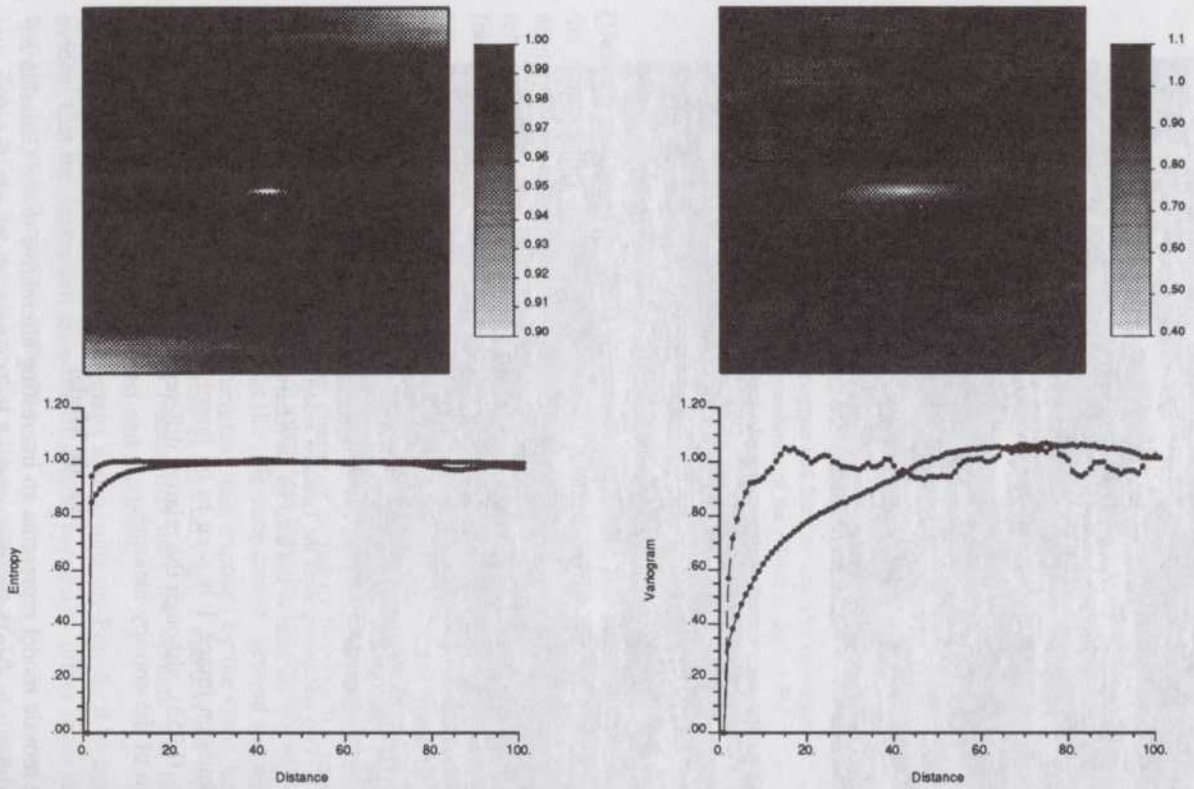

Fig. 6. Relative entropy and variogram of the first Gaussian realization. The upper left map is that of bivariate entropy, the lower left graph shows the vertical and horizontal cross-sections through the entropy map. The upper right map is that of the relative variogram, and the lower right graph shows cross-sections through the variogram map.

with common univariate Gaussian distribution. The first binormal distribution has a correlation coefficient of 1 (i.e., all values are equal) and the second a correlation coefficient of 0 (i.e., all values are independent). A mixture of these two bivariate Gaussian distributions with proportion $p(\mathbf{h})$ yields the following bivariate cumulative distribution function (Marechal, 1984):

$$\text{Prob}\{Y(\mathbf{u}) \leq z,\ Y(\mathbf{u}+\mathbf{h}) \leq \mathbf{z}'\} = p(\mathbf{h})\,G(\min(z, z')) + [1 - p(\mathbf{h})]\,G(z)\,G(z') \quad (16)$$

where $Y(\mathbf{u})$ represents the stationary RF model with standard normal cumulative distribution function (cdf): $G(y) = \text{Prob}\{Y(\mathbf{u}) \leq y\}$, $\forall \mathbf{u}$.

The previous bivariate distribution (16) corresponds to a RF model $Y(\mathbf{u})$ such that $Y(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$ are equal with probability $p(\mathbf{h})$ and independent with the complement probability $[1 - p(\mathbf{h})]$. It can be shown that the correlogram and all indicator correlograms of the RF $Y(\mathbf{u})$ are all equal to $p(\mathbf{h})$:

$$\frac{C_Y(\mathbf{h})}{C_Y(0)} = p(\mathbf{h}) \quad (17)$$

with $C_Y(\mathbf{h}) = \text{Cov}\{Y(\mathbf{u}), Y(\mathbf{u}+\mathbf{h})\}$

$$\frac{C_I(\mathbf{h}; y, y')}{C_I(0; y, y')} = p(\mathbf{h}),\ \forall y, y' \quad (18)$$

with $C_I(\mathbf{h}; y, y') = \text{Cov}\{I(\mathbf{u}; y), I(\mathbf{u}+\mathbf{h}; y')\}$ and $I(\mathbf{u}; y) = 1$ if $Y(\mathbf{u}) \leq y$; $= 0$ if not.

Thus, from its very definition, the mosaic model is expected to show less entropy (spatial disorder) than the Gaussian model with the same correlogram $p(\mathbf{h})$.

The mosaic RF $Y(\mathbf{u})$ is used to model the normal score transforms of the permeability reference data. The correlogram $p(\mathbf{h})$ is identified to the model $1 - \gamma(\mathbf{h})$ defined in (15).

Realizations of the mosaic model can be obtained by sequential indicator simulation, identifying all indicator correlograms to the unique correlogram $p(\mathbf{h})$ (see Eq. (18); Journel, 1989, p. 34). These realizations have a standard normal univariate cdf $G(y)$, hence must be backtransformed into simulated permeability values using the reference distribution of Fig. 2.

One hundred realizations of the mosaic model were generated using the program sisim of GSLIB (Deutsch and Journel, 1992); the first two generated realizations are given in Fig. 7. Once again, all realizations share the original histogram of Fig. 2. The relative entropy and variogram of the first realization are given on Fig. 8.

Figures 7 and 8 are to be compared to Figs. 1 and 3 for the reference and to Figs. 5 and 6 for the Gaussian model. Remember that all images share the

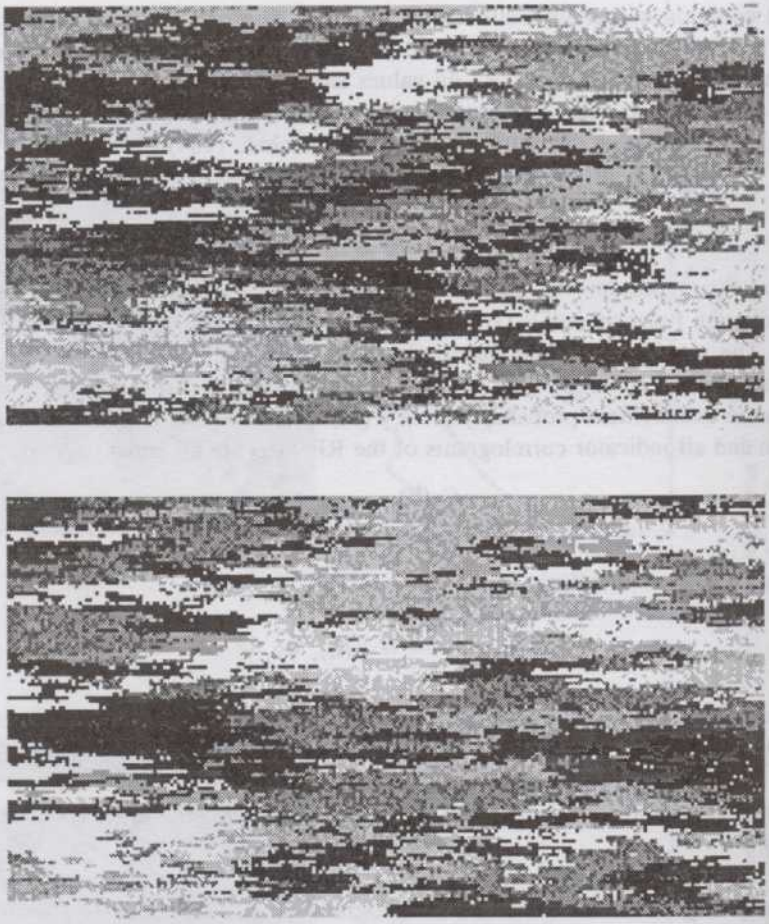

Fig. 7. Two realizations from the mosaic model. The unit scale, in millidarcies, is the same as that used for the reference image of Fig. 1.

same histogram, i.e., the same proportion of high/median/low permeability values, and the same variogram model (15) for normal score transforms. Hence, any difference is due to the implicit RF model beyond histogram and covariance.

The visual difference between the Gaussian and mosaic realizations is quite striking, with the mosaic realizations presenting clearer spatial structures (less entropy). From Table I, it appears that the mosaic model images have lower average bivariate relative entropy than the Gaussian model images.

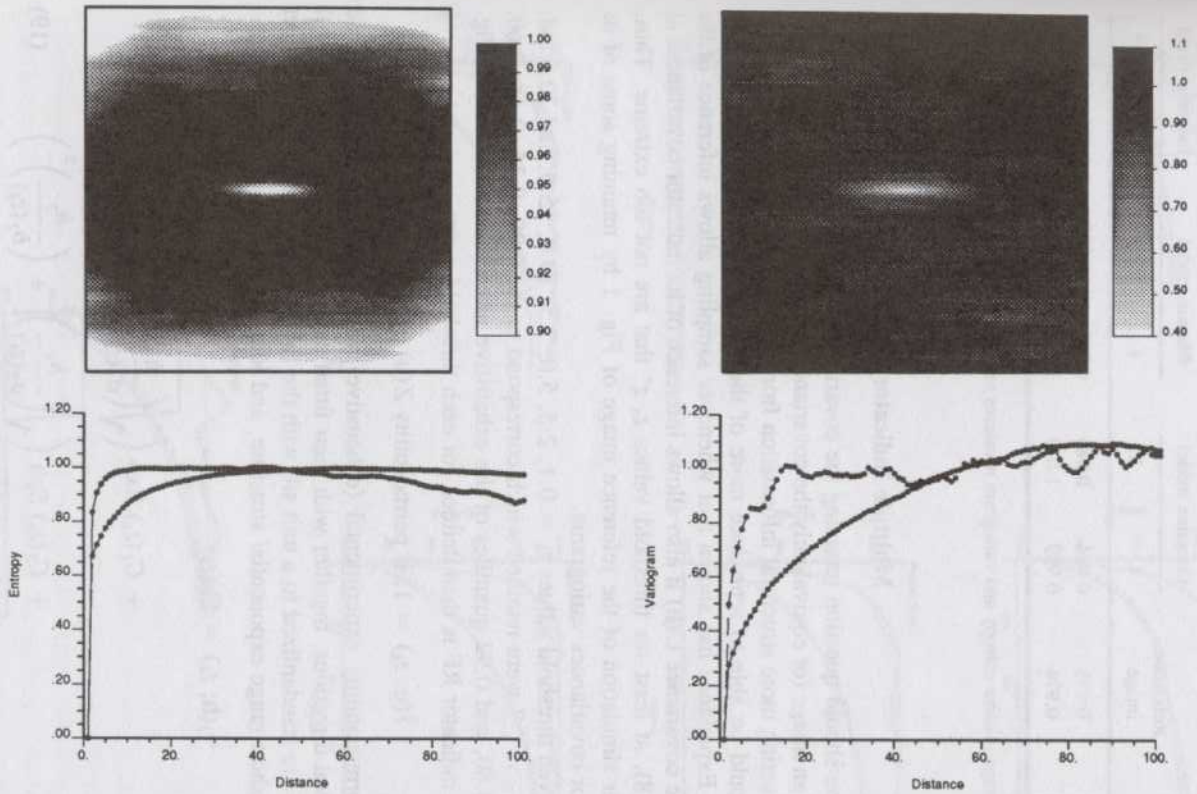

Fig. 8. Relative entropy and variogram of the first mosaic realization. The upper left map is that of bivariate entropy, the lower left graph shows the vertical and horizontal cross-sections through the entropy map. The upper right map is that of the relative variogram, and the lower right graph shows cross-sections through the variogram map.

Table I. Measures of Spatial Disorder[a]

	Reference image	Gaussian model		Mosaic model		Indicator model	
		1	2	1	2	1	2
$\overline{H_R}(A, A)$	0.953	0.994	0.985	0.962	0.938	0.928	0.913
$\overline{\gamma_R}(A, A)$	0.974	0.999	1.014	1.008	1.021	1.033	0.955

[a]The average relative entropy and variogram measures are defined over a square area A of side 100 pixels.

Multiple Indicator Model

One should question retaining the covariance as the sole spatial characteristic of an image (or equivalently the covariance of its normal scores transform). By extracting more structural information from the reference or sampled image, one should be able to reproduce more of the spatial features of the reference image. Experience has shown that whenever sampling allows inference of the attribute covariance $C(\mathbf{h})$ it also allows inference of its indicator covariances of type (18), at least for threshold values z, z' that are not too extreme. Thus, consider simulation of the reference image of Fig. 1 by retaining some of its indicator covariances/variograms.

Seven threshold values $z_k = 0.1, 2.5, 5.0, 7.5, 10.0, 15.0$, and 30.0 md, $k = 1, \ldots, 7$ were retained which correspond to the 0.20, 0.35, 0.48, 0.60, 0.67, 0.80, and 0.92 quantiles of the exhaustive reference distribution of Fig. 2. The indicator RF is then defined for each threshold z_k as:

$$I(\mathbf{u}; z_k) = 1, \text{ if permeability } Z(\mathbf{u}) \leq z_k; = 0 \text{ if not}$$

The corresponding experimental (exhaustive) variograms in the vertical and horizontal directions, together with their fitted models are given in Fig. 9. All models are standardized to a unit sill with the following combination of nugget effect, short range exponential structure, and longer range spherical structure:

$$\begin{aligned}\gamma_I(\mathbf{h}; z_k) = {} & C_0(z_k) \\ & + C_1(z_k)\,\mathrm{Exp}\left(\sqrt{\left(\frac{h_x}{a_1(z_k)}\right)^2 + \left(\frac{h_y}{b_1(z_k)}\right)^2}\right) \\ & + C_2(z_k)\,\mathrm{Sph}\left(\sqrt{\left(\frac{h_x}{a_2(z_k)}\right)^2 + \left(\frac{h_y}{b_2(z_k)}\right)^2}\right) \qquad (19)\end{aligned}$$

with Exp ( ) and Sph ( ) defined as in (15).

Table II gives the corresponding model parameters. Some remarks:

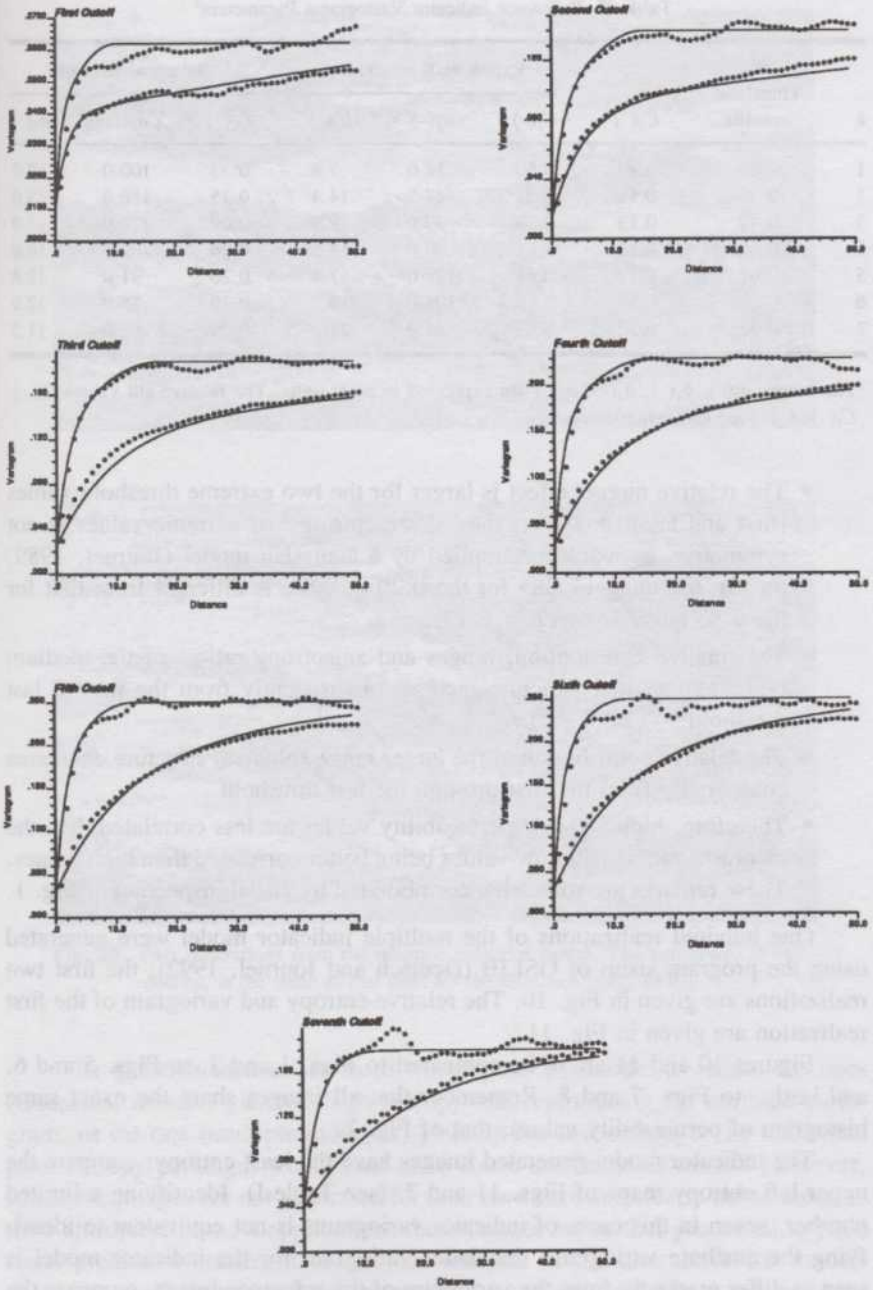

Fig. 9. Reference indicator variograms and their fitted models.

Table II. Reference Indicator Variograms Parameters[a]

			Exponential structure			Spherical structure		
k	Threshold quantile	$C_0(\ )$	$C_1(\ )$	$a_1(\ )$	$b_1(\ )$	$C_2(\ )$	$a_2(\ )$	$b_2(\ )$
1	0.20	0.17	0.50	18.0	7.8	0.33	100.0	10.0
2	0.35	0.11	0.54	47.7	14.4	0.35	150.0	15.0
3	0.48	0.13	0.58	90.0	18.0	0.29	170.0	17.0
4	0.60	0.13	0.61	90.0	14.5	0.26	160.0	16.0
5	0.67	0.12	0.68	108.0	17.4	0.20	91.0	12.8
6	0.80	0.12	0.68	108.0	15.2	0.20	85.0	12.0
7	0.92	0.22	0.69	144.0	21.5	0.09	66.0	11.2

[a]The ranges $a_1(\ )$, $b_1(\ )$, $a_2(\ )$, $b_2(\ )$ are expressed in pixel units. The relative sill values $C_0(\ )$, $C_1(\ )$, $C_2(\ )$ are dimensionless.

- The relative nugget effect is larger for the two extreme threshold values (first and last); however, the "destructuring" of extreme values is not symmetric, as would be implied by a Gaussian model (Journel, 1989, p. 33): the nugget effect for the 0.20 quantile is different from that for the 0.80 quantile ($0.17 \neq 0.12$).
- The relative contribution, ranges and anisotropy ratios, of the medium range exponential structure increases consistently from the first to last threshold.
- The relative contribution of the larger range spherical structure decreases consistently from the first through the last threshold.
- Therefore, high and low permeability values are less correlated than the medium values, with low values being better correlated than high values. These remarks are somewhat corroborated by visual inspection of Fig. 1.

One hundred realizations of the multiple indicator model were generated using the program sisim of GSLIB (Deutsch and Journel, 1992); the first two realizations are given in Fig. 10. The relative entropy and variogram of the first realization are given in Fig. 11.

Figures 10 and 11 are to be compared to Figs. 1 and 3, to Figs. 5 and 6, and lastly, to Figs. 7 and 8. Remember that all images share the exact same histogram of permeability values, that of Fig. 2.

The indicator model-generated images have the least entropy: compare the upper left entropy maps of Figs. 11 and 3, (see Table I). Identifying a limited number, seven in this case, of indicator variograms is not equivalent to identifying the attribute variogram: that latter variogram for the indicator model is seen to differ markedly from the variogram of the reference image, compare the right parts of Figs. 11 and 3.

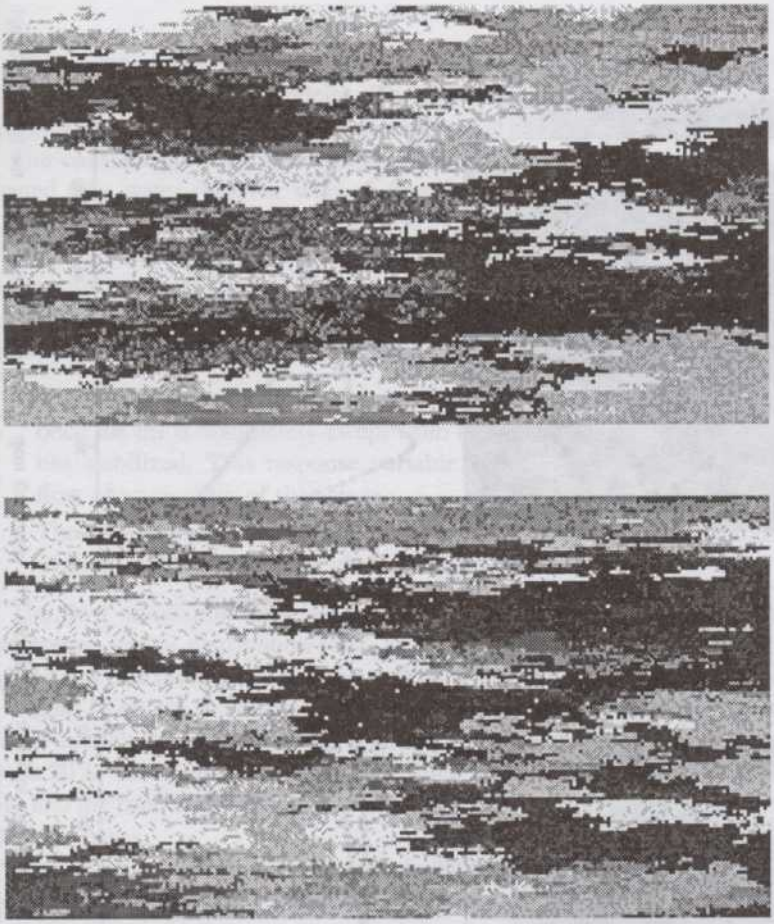

Fig. 10. Two realizations from the multiple indicator model. The unit scale, in millidarcies, is the same as that used for the reference image of Fig. 1.

The question then arises as to which prior RF model to retain, the ones (Gaussian, mosaic) yielding high entropy and reproducing the attribute variogram, or the one (multiple indicator) yielding too low entropy? The maximum prior entropy paradigm would lead us to chose the Gaussian model. However, stochastic images are not generated for their own sake and clearly not to measure their entropies; in most applications these images are further processed to yield some prediction statement. Thus, our analysis of the alternative RF models should go well beyond entropy, variogram measures, and visual inspection of the consequent stochastic images.

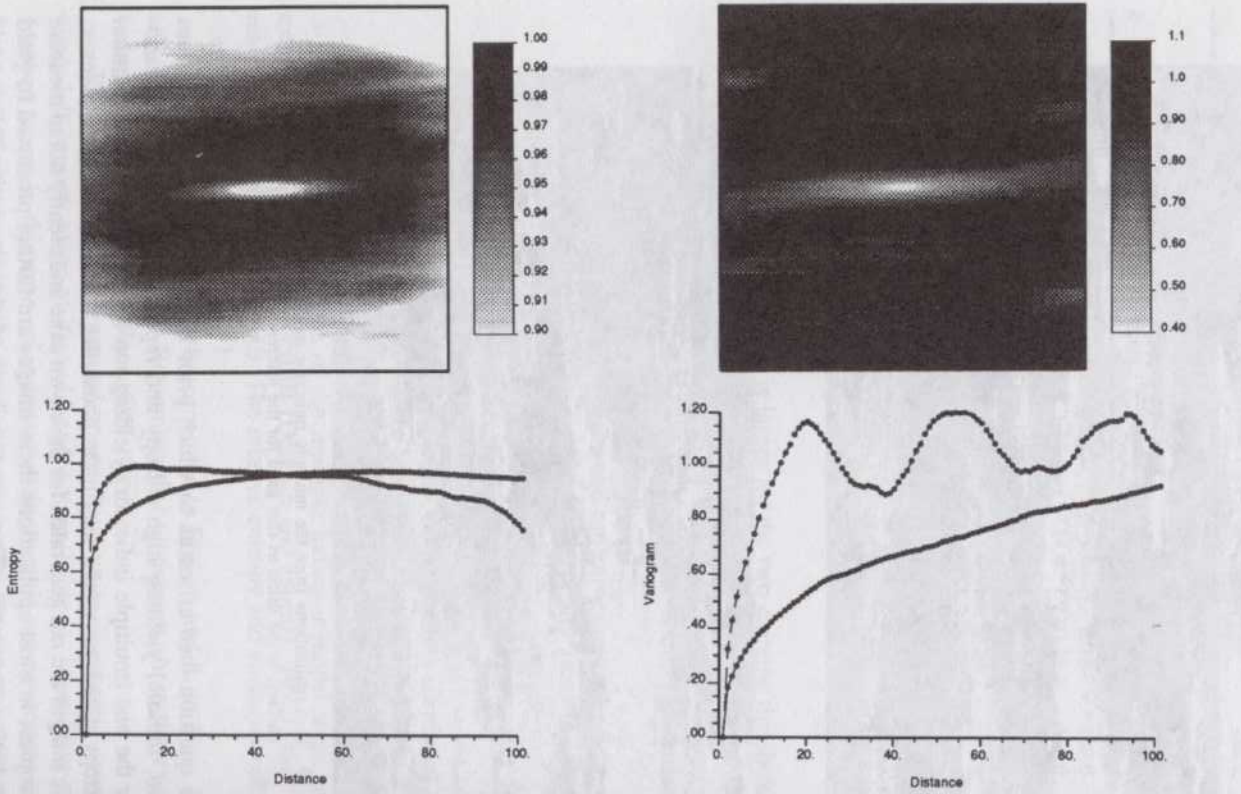

Fig. 11. Relative entropy and variogram of the first multiple indicator realization. The upper left map is that of bivariate entropy, the lower left graph shows the vertical and horizontal cross-sections through the entropy map. The upper right map is that of the relative variogram, and the lower right graph shows cross-sections through the variogram map.

Flow Modeling

Waterflood simulation was performed on all 300 images (100 realizations from each of the three RF models). The 2-D images were taken to represent vertical cross-sections between an injecting and a producing well.

The constant porosity grid blocks are initially saturated with oil and then subjected to a water flooding with no-flow boundary conditions on the upper and lower boundaries of the cross-section. The Eclipse flow simulator [1984] has been used for the flow simulations. Both wells operate at constant bottom hole pressure, straight line relative permeabilities were used, and both the oil and water have a unit mobility ratio. The following two flow response variables were isolated to characterize the flow characteristics of each image:

- The effective permeability of the image. This parameter is computed once the oil is completely swept from the image and the water flow rate has stabilized. This response variable provides an overall steady state flow characteristic of the image.
- The time at which the water cut reaches 90%. This response is a measure of sweepage and late flow characteristics of the image.

The flow simulation exercise described above was also carried out using the reference spatial distribution (Fig. 1). The reference effective permeability was 8.33 md and the time to achieve a 90% water cut was 3.70 time units.

Figure 12 shows the histograms of effective permeability obtained from the three RF models being considered. Figure 13 shows the histograms of the time to reach 90% water cut obtained from the three RF models. The effective permeability and the time to achieve 90% water cut have a correlation of about -0.75 in all three cases.

Table III gives the mean, standard deviation (σ), minimum and maximum value of the flow response variables for the three sets of simulations.

Observations

- The actual reference time value (3.70 time units) is not within the range of the 100 Gaussian model-derived response values and barely within the ranges of the mosaic and indicator model-derived distributions. The reference effective permeability (8.33 md) is only barely in the range of the Gaussian model derived-distribution and fits only slightly better in the ranges of the mosaic and indicator model-derived distributions.
- Although all 300 realizations of the permeability field reproduce *exactly* the reference univariate distribution (Fig. 1) they all yield distributions which centers (mean or median) deviate considerably from the reference value. Moreover, the Gaussian model is inaccurate in the sense that it

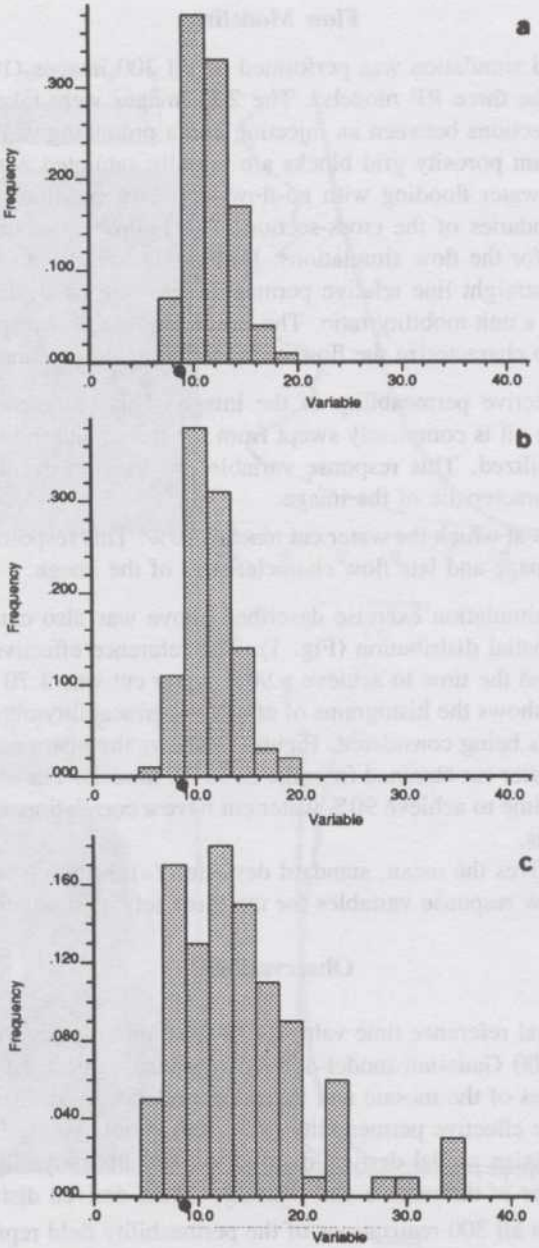

Fig. 12. Distribution of effective permeabilities obtained from the (a) Gaussian, (b) mosaic, and (c) indicator RF models. The dot location gives the reference image value (8.33 md).

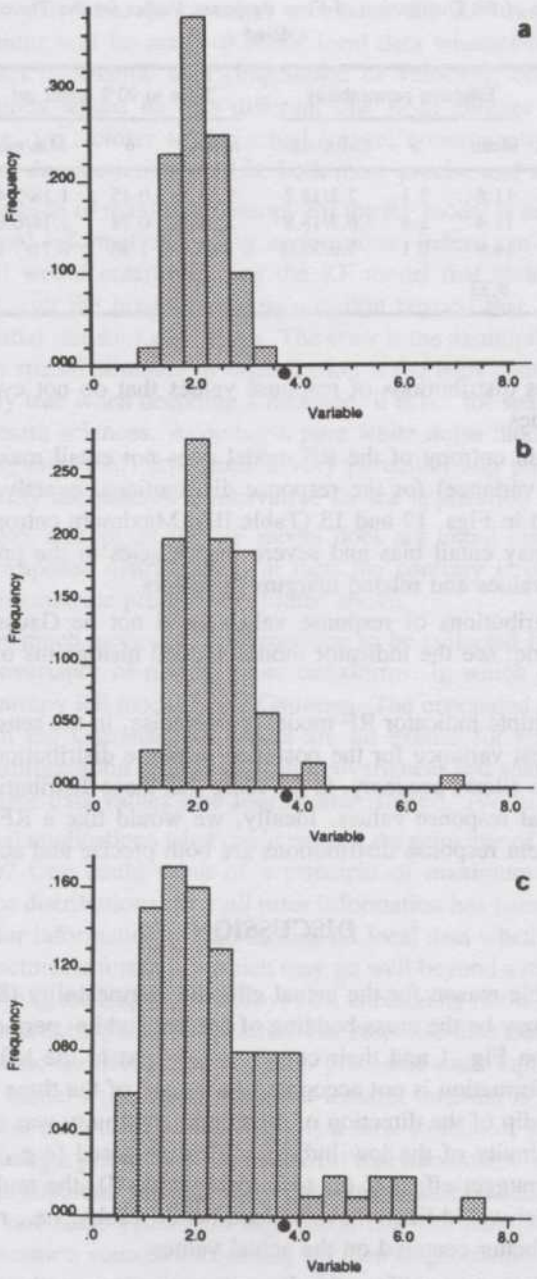

Fig. 13. Distribution of late breakthrough times obtained from the (a) Gaussian, (b) mosaic, and (c) indicator RF models. The dot location gives the reference image value (3.70 time units).

Table III. Statistics of the Distribution of Flow Response Values for the Three RF Models Considered

	Effective permeability			Time to 90% water cut			
RF model	Mean	σ	Min/max	Mean	σ	Min/max	Entropy (*)
Gaussian	11.8	2.1	7.2/18.2	2.11	0.45	1.18/3.48	0.994
Mosaic	11.4	2.4	6.3/18.8	2.34	0.78	1.14/6.81	0.962
Indicator	14.0	6.1	5.0/35.0	2.41	1.30	0.71/7.15	0.928
Reference inage	8.33			3.70			0.953

generates distributions of response values that do not even contain the true values.

- Maximum entropy of the RF model does not entail maximum entropy (largest variance) for the response distributions: exactly the reverse is observed in Figs. 12 and 13 (Table III). Maximum entropy of the input model may entail bias and severe inaccuracies in the prediction of response values and related margins for errors.
- The distributions of response values need not be Gaussian, nor even symmetric: see the indicator model-derived histograms of Figs. 12 and 13.
- The multiple indicator RF model is imprecise, in the sense that it yields the largest variance for the posterior response distribution; but that imprecision allows accuracy, in the sense that these distributions do include the actual response values. Ideally, we would like a RF model whose consequent response distributions are both precise and accurate.

DISCUSSION

- A possible reason for the actual effective permeability (8.33 md) to be so low may be the cross-bedding of the low (white) permeability streaks visible on Fig. 1 and their concentration next to the left vertical well. Such information is not accounted for in any of the three RF models. If the 10° dip of the direction of maximum continuity was considered and the continuity of the low indicator RVs increased (e.g., by setting the relative nugget effect $C_0(1)$ to zero in Table II), the multiple indicator RF model would likely have yielded better results, i.e., response distributions better centered on the actual values.
- Purposely not to confuse the discussion on entropy, the stochastic simulations were made nonconditional to local data, e.g., the producer and

injector well data. In practice, no matter the RF model chosen, the realizations will be made to honor local data whether originating from well logs or seismic data (impedance or velocity); consequently, the realizations would be less different one from another and also more accurate, i.e., closer to the actual image; consequently, the posterior response distributions would be both more precise and more accurate.

- The principle of maximum entropy for the RF model is debatable beyond traditional information coding applications. Indeed, in spatial applications, it would entail choosing the RF model that yields images/realizations with the least spatial organization beyond that imposed by the data spatial statistics and values. The error is the assumption that a model with no spatial structure or organization is the least committing. This is certainly true when decoding a message, it is not for spatial applications in the earth sciences. Adopting a pure white noise model (the ultimate in disorganization!) represents a very particular and precise model not necessarily adequate or conservative for the application intended.
- Maximum entropy of the RF model does not entail maximum entropy of the response distributions; in fact, the contrary is observed for the flow performance predictions studied above.
- There is much more actual information to be collected from data than a mere covariance of normal score transforms. In which case, the maximum entropy RF model is not Gaussian. The unequaled analytical properties of the Gaussian RF model are no excuse for ignoring possibly critical information such as indicator covariances and spatial connectivity of extreme data values (see Journel and Alabert, 1989).
- In spatial applications what could replace the principle of maximum prior entropy? One could think of a principle of maximum entropy of the response distributions once *all* prior information has been accounted for. The prior information should include all local data whether hard or soft, and structural information which may go well beyond a mere covariance. Maximizing the response entropy, i.e., uncertainty of the response values is necessarily application-specific. The response distributions depend on the specific transfer function which processes each input realization of the RF model $Z(\mathbf{u})$. Most often that transfer function is highly complex and is not analytically defined, e.g., a flow simulator; then, the task of determining a priori the RF model $Z(\mathbf{u})$ that maximizes response entropy becomes hopeless. Instead, one should concentrate on building RF models and/or stochastic simulation algorithms that can handle a large variety of information sources accounting for their imprecision. Note that more prior information does not necessarily entail a narrower (more precise) distribution of response values because the RF model is likely to change

as more information is considered. Geologists know that the true complexity of a field is gradually revealed as more information becomes available.

CONCLUSION

The property of maximum entropy of the Gaussian RF model is conditional to retaining a single covariance function from the data. In the earth sciences, the most interesting processes are the ones that present characteristic spatial structures that cannot be described by a single covariance function; rather, geometric shapes and a series of indicator covariances characterizing spatial connectivity would be needed. Then, the maximum entropy RF model is no longer Gaussian.

In any case, maximum entropy of the RF model distribution does not entail maximum entropy of the response distributions obtained from processing realizations of that RF model. For some applications, using a maximum entropy RF model may lead to inaccurate prediction of response values and an overoptimistic (narrow) assessment of uncertainty of such prediction. Too narrow probability intervals may not contain the "true" value.

The entropy to be maximized is that of the response (output) distributions, not that of the input realizations of the RF model. A pure white noise model may be maximum entropy; however, when processed through flow simulators it may yield very low entropy response distributions both inaccurate and imprecise.

It is suggested that research effort be dedicated to building flexible families of RF models that can accommodate a large variety of prior information other than a mere covariance and, in particular, fuzzy and uncertain data. Also the time to randomize global statistics such as the histogram and covariance is long overdue.

ACKNOWLEDGMENTS

Financial support for this research was provided by the industrial affiliate member companies of the Stanford Center for Reservoir Forecasting (SCRF).

REFERENCES

Christakos, G., 1990, A Bayesian/Maximum-Entropy View to the Spatial Estimation Problem: Math Geol., v. 22, n. 7, p. 763–777.

Deutsch, C., and Journel A., 1992, GSLIB: Geostatistical Software Library and User's Guide: Oxford University Press, New York.

ECLIPSE Reference Manual: Exploration Consultants Limited, Highlands Farm, Greys Road, Henley-on-Thames, Oxon, England, November 1984.

Jaynes, E., 1985, Entropy and Search Theory, *in* C. Smith and W. Gandy, Jr. (Eds.), Maximum Entropy and Bayesian Methods in Inverse Problems: Reidel, Dordrecht, Holland, p. 443–454.

Jaynes, E., 1985, Where Do We Go from Here? *in* C. Smith and W. Gandy, Jr. (Eds.), Maximum Entropy and Bayesian Methods in Inverse Problems: Reidel, Dordrecht, Holland, p. 21–58.

Jones, D., 1979, Elementary Information Theory: Clarendon Press, Oxford.

Journel, A., 1983, Non-Parametric Estimation of Spatial Distributions: Math. Geol., v. 15, n. 3, p. 445–468.

Journel, A., 1989, Fundamentals of Geostatistics in Five Lessons, Vol. 8. Short Course in Geology: American Geophysical Union, Washington, D. C.

Journel, A., and Alabert, F., 1989, Non-Gaussian Data Expansion in the Earth Sciences: Terra Nova, v. 1, p. 123–134.

Journel A., and Huijbregts, C. J., 1978, Mining Geostatistics: Academic Press, New York.

Kullback, S., 1968, Information Theory and Statistics: Dover, New York.

Marechal, A., 1984, Recovery Estimation: A Review of Models and Methods, *in* G. Verly et al. (Eds.), Geostatistics for Natural Resources Characterization: Reidel, Dordrecht, Holland, p. 385–420.

Scholle, P., and Spearing, D. (Eds.), 1982, Sandstone Depositional Environments: The American Association of Petroleum Geologists, Tulsa, Oklahoma.

Shannon, C., 1984, A Mathematical Theory of Communication: Bell System Tech. J., v. 27, p. 379–623.