Scaling Multiple Point Statistics to Different Histograms

Steven Lyster[1], Julián M. Ortiz[2], and Clayton V. Deutsch[1]

[1]Department of Civil & Environmental Engineering
University of Alberta

[2]Departamento de Ingeniería de Minas
Universidad de Chile

Abstract

Multiple-point statistics can be used to model high-order relations found in complex geologic features. These statistics are inferred from a training image deemed representative of the area to be simulated. The training image must contain all the relevant geologic features and structures. Often, a representative training image will have the wrong histogram or single-point distribution. A large scale trend causes the histogram to vary locally. For these reasons, an effective and simple method for scaling multiple-point histograms to reflect changes in the single-point distribution is desirable. This paper presents a robust method to scale/change multiple-point statistics to reflect arbitrary changes in the histogram. The method works by applying an equation to each class in the multiple-point histogram to account for the change in every facies type. The equation also accounts for the different proportions of facies contained in each class.

Introduction

Simulation of geological heterogeneity using geostatistical methods is an increasingly common practice. The variogram is the typical statistic used to express spatial continuity. Variograms adequately model the two-point connectivity and are simple to infer from data. Robust simulation techniques and software are have been developed for simulation with a variogram; however, many geologic phenomena do not follow linear patterns and cannot be properly modeled relying only on variograms. Curvilinear and higher-order arrangements are common. Multiple-point statistics can be used to better express these complex arrangements.

Several different methods have been explored for utilizing multiple-point statistics in simulations. Using extended normal equations [4,7] is a direct method for honouring multiple-points statistics. Neural networks [1] utilize an iterative approach to enforce the desired distribution. Simulated annealing [2] perturbs an image based on an objective function and the Boltzmann probability distribution. All of these techniques infer the multiple-point statistics from a *training image* that must be exhaustively defined at the scale of the simulated grid and represents the area being simulated. Establishing a training image is not straightforward. I t must be at the scale and resolution we wish to simulate. We could not draw such an image because we would be unable to successfully position the fine scale details and structure of natural heterogeneity. Outcrop images may provide some assistance, but the orientation and irregularity of the outcrop would make it impossible to infer a 3-D image at the required resolution. We often resort to process- or object-based simulation techniques to create a training image.

The inference and utilization of multiple-point statistics can be difficult. A large amount of data is required to infer even simple statistics adequately. The assumption of stationarity calls for extra

care to be taken in the determination of the multiple-point statistics. Because multiple-point statistics contain all lower-order statistics within the same configuration, even slight variations in the single-point distributions can have significant impact on high-order relations. Adopting the multiple-point statistics from a training image amounts to adopting the histogram that goes with it.

The multiple-point histogram could be scaled to match different single-point distributions for different areas. The multiple-point statistics could be inferred from a single training image, and then scaled to varying single-point histograms for different zones. This would be useful in cases where stationarity can not be assumed.

Scaling a multiple-point histogram could also be used to correct a bias in the training image; for example, if an outcrop map correctly identifies the underlying geologic structures but does not properly represent the proportions of facies for an area to be simulated.

Multiple-Point Statistics

Two-point statistics often cannot fully characterize geological arrangements containing curvilinear features and other complex shapes. Variograms have difficulty in characterizing distinct, complex shapes. An example of this is shown in Figure 1: with two facies the variogram for each will be the same. Indicator simulation cannot differentiate between black spheres on a white background and white spheres on a black background. Multiple-point simulation can differentiate between the facies.

The most common multiple-point statistic to be considered is the multiple-point histogram [3]. Continuous variables are divided by cutoffs into $k=1,\ldots,K$ classes so every point is easily assigned one class. Facies for categorical variables are numbered in a similar way. Then, N points in a given arrangement are considered. For a given statistic there are K^N classes. If the points are numbered $n=1,\ldots,N$, the index j of each class can be found by the equation

$$j = 1 + \sum_{n=1}^{N}\bigl(k(n)-1\bigr)\times K^{n-1} \tag{1}$$

This assigns a unique index j to every possible arrangement in the statistic [2].

Multiple-point statistics also contain all lower-order statistics as well. It is quite easy to calculate the single-point distribution from the multiple-point histogram. If the single-point proportion of facies k is represented by P_k then it can be found using f_j and $P_{k,j}$, which is the proportion of facies k in class j. $P_{k,j}$ can be found by converting $(j - 1)$ to a base K number in the following process[1]:

      1. Take $j - 1$ and divide it by K

      2. The remainder of this division represents point $n = 1$.

      3. Divide the quotient from step 1 by K again.

      4. The remainder of this division represents another point.

      5. Repeat 3 and 4 until N points have been represented.

      6. To get $P_{k,j}$ for facies k, divide the number of points $(k - 1)$ by N.

---

[1] The method for converting a number to a different base was found online at
http://mathbits.com/MathBits/CompSci/Introduction/frombase10.htm

With all the f_j and $P_{k,j}$ values, the single-point proportions can be found using the equation:

$$P_k = \sum_{j=1}^{K^N} P_{k,j} \times f_j \qquad (2)$$

An example is shown in Table 1. The statistic used is four points, with two facies in the training image. Other lower-order statistics could be found using a similar method.

Class, j	Point 1	Point 2	Point 3	Point 4	$P_{1,j}$	$P_{2,j}$	Frequency, f_j	$P_{1,j} \times f_j$	$P_{2,j} \times f_j$
1	1	1	1	1	1	0	0.0081	0.0081	0
2	2	1	1	1	0.75	0.25	0.0189	0.014175	0.004725
3	1	2	1	1	0.75	0.25	0.0189	0.014175	0.004725
4	2	2	1	1	0.5	0.5	0.0441	0.02205	0.02205
5	1	1	2	1	0.75	0.25	0.0189	0.014175	0.004725
6	2	1	2	1	0.5	0.5	0.0441	0.02205	0.02205
7	1	2	2	1	0.5	0.5	0.0441	0.02205	0.02205
8	2	2	2	1	0.25	0.75	0.1029	0.025725	0.077175
9	1	1	1	2	0.75	0.25	0.0189	0.014175	0.004725
10	2	1	1	2	0.5	0.5	0.0441	0.02205	0.02205
11	1	2	1	2	0.5	0.5	0.0441	0.02205	0.02205
12	2	2	1	2	0.25	0.75	0.1029	0.025725	0.077175
13	1	1	2	2	0.5	0.5	0.0441	0.02205	0.02205
14	2	1	2	2	0.25	0.75	0.1029	0.025725	0.077175
15	1	2	2	2	0.25	0.75	0.1029	0.025725	0.077175
16	2	2	2	2	0	1	0.2401	0	0.2401
							Σ	0.3	0.7

Table 1: An example showing how Equation (2) can be used to derive a single-point histogram from a multiple-point one. The class frequencies were inferred from a simple training image. The multiple-point histogram has $N=4$ points and there are $K=2$ facies in the image.

Scaling the Histogram

Most simulation techniques leave some room for variance from the target histogram; however, there is still dependence on the training image and the assumption of stationarity. It is possible that the training image captures the proper essence of the area to be simulated, without correctly matching the desired single-point distribution. This could be the case for locally varying single-point distributions where the geologic structures are global. An example of this is in Figure 2: the four pictures show realizations with the same structures but different proportions of the facies.

Scaling would allow differing proportions of facies in several defined zones without the need to develop multiple training images. It would be necessary to preserve the structures of the facies while easily changing the proportions. A simple equation would be the best, but could be difficult to determine. There are also many possible equations that could be used to present scaled histograms; determining the "best" one would be very subjective.

Random Case

The simplest case for determining a multiple-point histogram is a totally random case. For uncorrelated points, the arrangement does not affect the class frequencies and no conditional probabilities are required. The frequency f_j of any class for an N-point statistic with K facies having single-point proportions P_k, $k=1,\ldots,K$ can be calculated by:

$$f_j = \prod_{k=1}^{K} P_k^{N \times P_{k,j}} \qquad (3)$$

Since $P_{k,j}$ is the proportion of facies k in class j, $N \times P_{k,j}$ is the number of occurrences of facies k in class j.

Scaling the random case is quite simple; multiplying the original frequency by a series product of the desired frequency over the old frequency results in a simple equation:

$$\hat{f}_j = f_j \times \prod_{k=1}^{K} \left(\frac{\hat{P}_k}{P_k} \right)^{N \times P_{k,j}} \qquad (4)$$

$$= f_j \times \frac{\prod_{k=1}^{K} \hat{P}_k^{N \times P_{k,j}}}{\prod_{k=1}^{K} P_k^{N \times P_{k,j}}}$$

$$= f_j \times \frac{\hat{f}_j}{f_j} = \hat{f}_j$$

where $\hat{P}_k$ is the target single-point proportion of facies k and $\hat{f}_j$ is the resulting frequency of class j. Figure 3 shows a randomly-populated field and its corresponding f_j values before and after scaling; the post-scaling frequencies perfectly match those expected both from Equation (3) and from a different randomly-populated field. For a non-random case, an extension of Equation (4) could lead to a more general formula with some modifications to account for correlation between the points.

Non-Random Case

For non-random cases Equation (4) returns P_k values that are scaled too much, that is, the probability classes are exaggerated. Classes that should be raised are raised beyond where they should be, and those that are lowered are lowered too much. The desired single-point statistics are not returned. Figure 4 shows a striped image and its calculated f_j values before and after scaling. Note that the scaled frequencies are obviously out of proportion as they sum to 1.21 instead of 1.

In order for a scaling technique to be useful and robust, it must return the desired P_k values and force the f_j values to sum to 1. Equation (4) does not properly do this for non-random images. Removing the N term from the exponent in Equation (4) causes the class frequencies to be under-

rather than over- scaled. This modification still scales the multiple-point classes proportionally based on which facies they contain; this is due to the $P_{k,j}$ in the exponent.

Removing the N term from the exponent in Equation (4) does not force the frequencies to sum to exactly 1; however, the frequencies are all scaled so that the single-point proportions approach the target values. This modified equation is therefore:

$$\hat{f}_j = f_j \times \prod_{k=1}^{K}\left(\frac{\hat{P}_k}{P_k}\right)^{P_{k,j}} \qquad (5)$$

This equation can be repeated, obtaining a new set of f_j and P_k values each time. Iterating Equation (5) leads to convergence with the frequencies summing to 1 and the single-point proportions precisely matching the desired distribution. The number of iterations performed does not need to be large, as convergence is approached rapidly. As the P_k values move closer to the target proportions the changes caused by each iteration become smaller.

This iterative approach works for the previously solved random case as well; after 50 iterations, all classes converge to within 0.0001% of their true calculated values from Figure 3. This amount of divergence is easily within the margin for error of any training image that would be developed. Performing 50 iterations requires less than a second on an old PC. The process is not memory intensive.

The random case takes longer to converge than others. This is because the scaling product values are the most extreme due to the significantly higher exponent in Equation (4) as compared to Equation (5). Better-ordered images require fewer iterations.

Example

Figure 5 shows training images created by truncated sequential gaussian simulation with 30% high values in the image at the top, with 35% and 50% high values in the lower left images. The high values are represented by the shaded areas. The different shades of grey represent different ranges of the high values, with darker being higher. In all the training images 20% of the high values are very high (assigned an arbitrary value of 3), 60% is somewhat high (value of 2), and 20% is low-high: lower than the rest but still above the background low values (value of 1). All of the low value points, shown as white, were assigned a value of 0. The lower right of Figure 5 shows two realizations produced by simulated annealing using a four-point statistic arranged in a 2x2 square pattern. One of the images has a target of 35% high values, the other 50%. With three high value codes and one low value code there are $4^4=64$ potential classes. All of the grids are 200x200 squares in size.

The simulations were performed using histograms calculated from the 30% high value training image then scaled iteratively with Equation (5) to match the proper distribution of high and low values. Scaling from 30% high value rock codes to 50% is probably at the limits of what would be done in practice; the 30% to 35% scaling is likely to be more common. The 50% scaling was done to show more variation in the change of support graphs for various sizes and to investigate the limits of practicality for scaling. Convergence for scaling the histogram was taken to be when the sum of the differences between the P_k values and the target P_k values was less than 10^{-7}. At this threshold, the difference between the target and scaled P_k values is insignificant on a grid with 40 000 nodes.

Upon visual inspection, the simulations do seem to capture the general arrangement of the facies properly; the very high values are in clusters of various sizes, surrounded by medium high values with low-high values on the outskirts. Some noise is obviously present in the background areas of the simulations.

Block averaging was performed to investigate simple response characteristics for the simulations and to give some quantification to the effectiveness of the technique used. Figure 6 shows the block averaged maps of the same images as Figure 5. The block averaging was performed over a 5x5 grid. Once again, visually the simulations appear to have similar form to the training images.

To check the results, cutoff grade-tonnage curves were made for all of the images. For the training images these were taken over 10 SGS realizations; for the multiple-point annealing simulations they were taken over 5 realizations. The resulting cutoff value vs. fraction above cutoff curves are shown in Figure 7. The curves for the scaled realizations match the training image curves well. Even the realizations using a histogram scaled from 30% high values to 50% high values is very close in shape to the reference curve. Some noise is present, but there is no systematic bias.

Procedures

Figure 8 shows a flowchart of the procedure used to calculate a multiple-point histogram. This can be used for any arrangement of N points for any size of field and number of facies, K. The basis behind the process is to go through all of the points in the training image field, then calculate what class j the statistic using that point as $n=1$ falls into. The number of occurrences of class j are counted and then divided by the total number of points for which an arrangement of n points can be found. This results in frequencies f_j for each unique K^N class in the statistic.

The flowchart shown in Figure 9 is the iterative scaling process proposed by this paper. It involves iterating Equation (5) as many times as needed to meet the convergence criteria. For the example given here convergence was taken when the sum of the differences between all target and scaled single-point proportions P_k was less than 10^{-7}. Other criteria could be used, such as a sum of squared differences.

Conclusions

A process to scale a multiple-point histogram to an arbitrary histogram is a valuable tool in the development of multiple-point statistics for simulation. The iterative approach is one possibility that could be used. The advantage of this approach is that it is unbiased and is independent of complex user inputs. Also, it correctly matches the target single-point distribution to whatever precision is desired. The disadvantage to a formula-based approach is the potential for oversimplification resulting in important features not being scaled to the desired degree and the essence of the training image not being properly retained. For small changes it is unlikely that iterating Equation (5) would result in errors in the histogram greater than the margin for error of the original training image.

References

1. J. Caers, A.G. Journel. Stochastic Reservoir Simulation Using Neural Networks Trained on Outcrop Data. SPE 49026. In *Proceedings of SPE ATCE*, New Orleans, LA, October 1998.

2. C.V. Deutsch. *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data*. PhD thesis, Stanford University, Stanford, CA, 1992.

3. C.V. Deutsch, E. Gringarten. Accounting for multiple-point continuity in geostatistical modeling. In *6th International Geostatistics Congress*, Cape Town, South Africa, April 2000. Geostatistics Association of Southern Africa.

4. F.B. Guardiano, R.M. Srivasava. Multivariate geostatistics: Beyond bivariate moments. In A. Soares, editor, *Geostatistics Troia '92*, volume 1, pages 133-144. Kluwer, 1993.

5. J.M. Ortiz. Characterization of High Order Correlation for Enhanced Indicator Simulation. PhD thesis, University of Alberta, Edmonton, AB, 2003.

6. J.M. Ortiz. Selected aspects of multiple-point statistics. In *Centre for Computational Geostatistics*, volume 5, Edmonton, AB, 2003.

7. S. Strebelle, A.G. Journel. Sequential simulation drawing structures from training images. In *6th International Geostatistics Congress*, Cape Town, South Africa, April 2000. Geostatistics Association of Southern Africa.

8. S. Strebelle, A.G. Journel. Reservoir Modeling Using Multiple-Point Statistics. SPE 71324. In *Proceedings of SPE ATCE*, New Orleans, LA, October 2001.
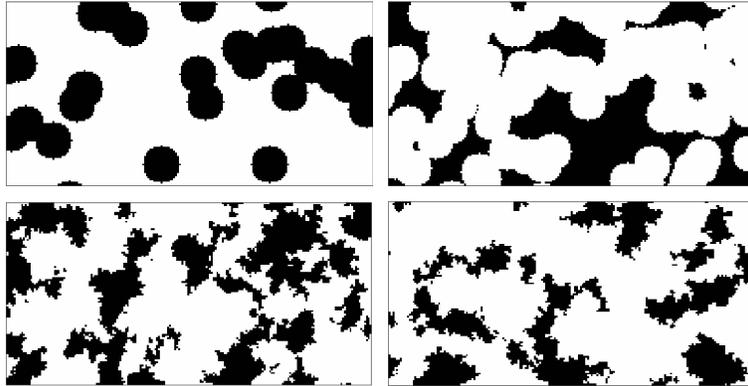
Figure 1: Top: Two simulated realizations with spheres as the geologic features. Each has 30% black area and 70% white area. Bottom: Sequential indicator simulated images produced using the variograms of the top images. Note that the SIS images appear the same; this is because of the inability of the variogram to distinguish between the top two images.
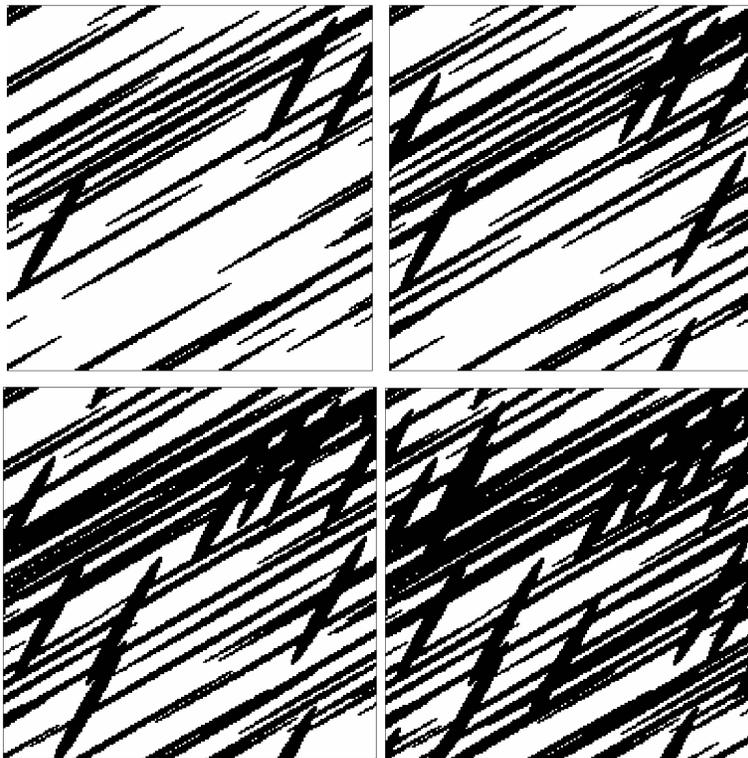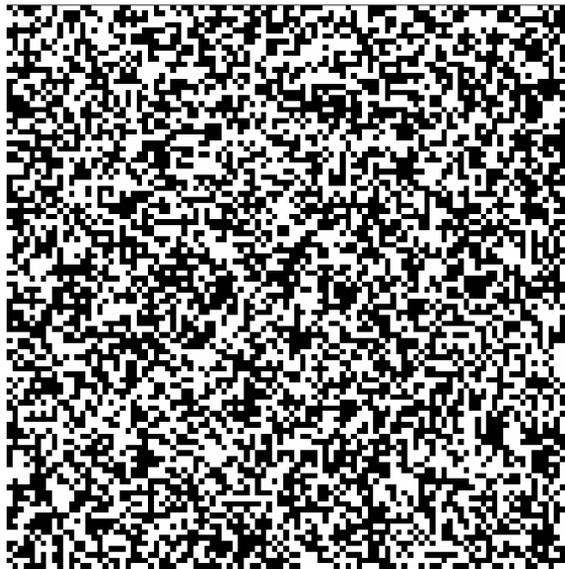

Figure 2: Four realizations of facies. The black ellipses in the top left, top right, bottom left, and bottom right make up 30%, 40%, 50%, and 60% of the fields respectively. The underlying features are the same in all realizations.

Class	f_j	$\hat{f}_j$
1	0.0625	0.1296
2	0.0625	0.0864
3	0.0625	0.0864
4	0.0625	0.0576
5	0.0625	0.0864
6	0.0625	0.0576
7	0.0625	0.0576
8	0.0625	0.0384
9	0.0625	0.0864
10	0.0625	0.0576
11	0.0625	0.0576
12	0.0625	0.0384
13	0.0625	0.0576
14	0.0625	0.0384
15	0.0625	0.0384
16	0.0625	0.0256

Figure 3: Frequencies pre- and post-scaled for a random image; $K=2$ and $N=4$. The proportions are scaled from $P_1=0.5$, $P_2=0.5$ to $P_1=0.6$, $P_2=0.4$.

Class	f_j	$\hat{f}_j$
1	0.45	0.93312
2	0	0
3	0	0
4	0.05	0.04608
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0.05	0.04608
14	0	0
15	0	0
16	0.45	0.18432

Figure 4: An image with stripes 10 units thick, $K=2$. The frequencies for a statistic with $N=4$ are to the right. The initial proportions are $P_1=0.5$, $P_2=0.5$; the target scaled proportions are $P_1=0.6$, $P_2=0.4$. The output proportions are $P_1=0.9792$, $P_2=0.2304$.

Training Image: 30% High values

Training Image: 35% High values

Simulated Image: 35% High values

Training Image: 50% High values
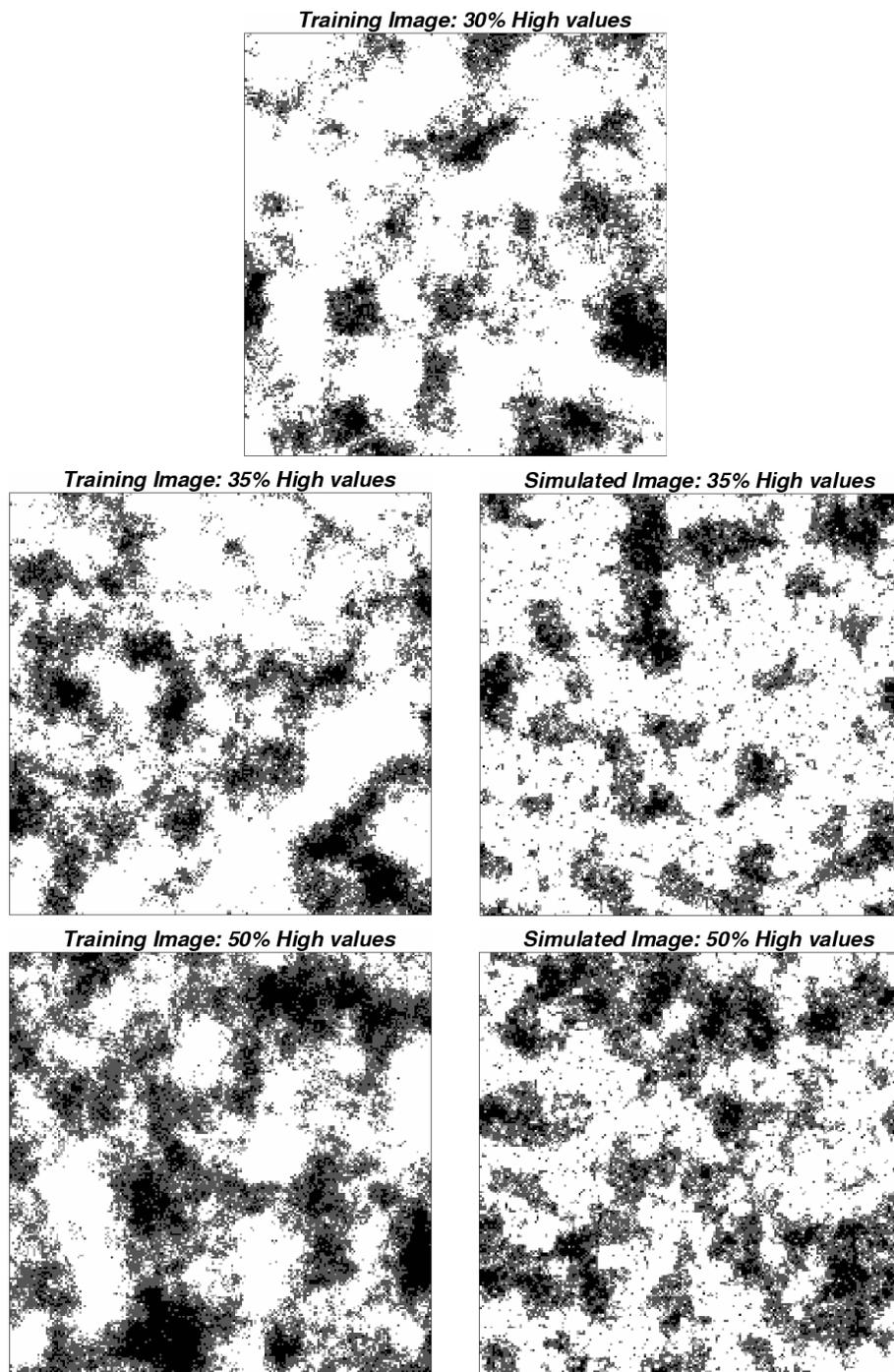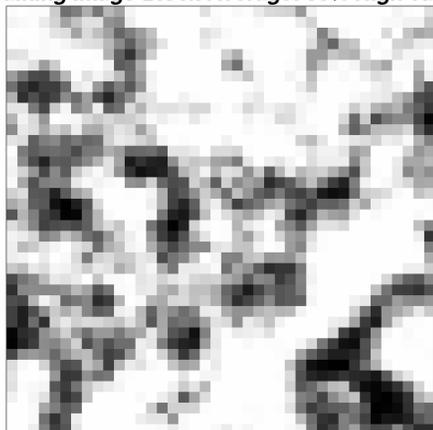
Simulated Image: 50% High values

Figure 5: Top: Truncated gaussian training image with 30% high values. Lower left: truncated gaussian training images with 35% and 50% high values. Lower right: simulated images created by multiple-point simulated annealing using histograms scaled from the 30% training image. Darker shading represents higher values.

Training Image Block Average: 30% High values

Training Image Block Average: 35% High values

Simulated Block Average: 35% High values

Training Image Block Average: 50% High values
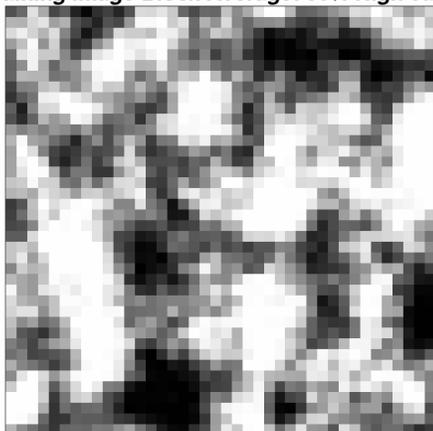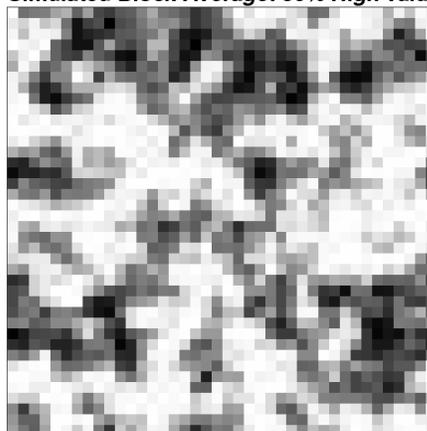

Simulated Block Average: 50% High values

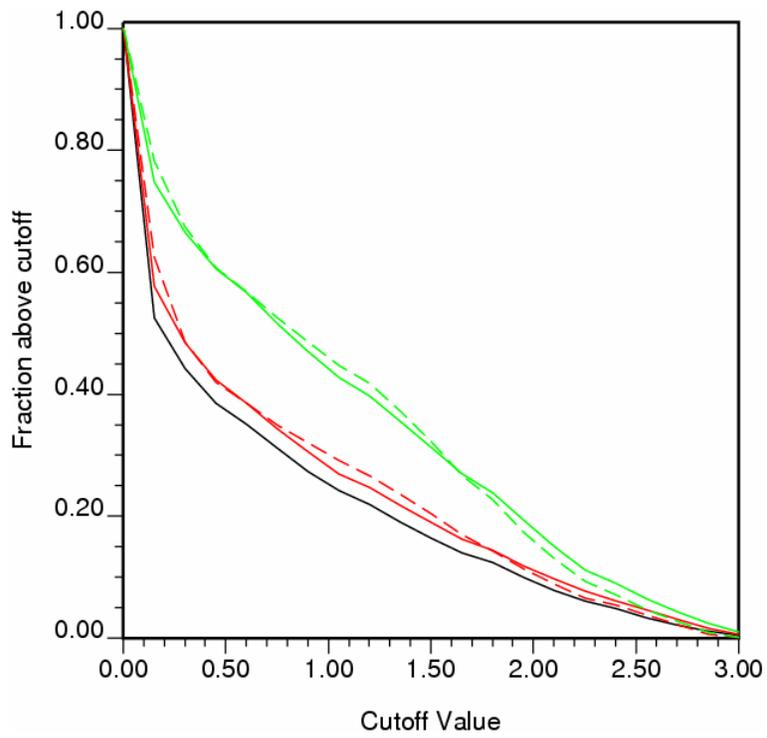**Figure 6**: 5x5 block averages of all the images from Figure 5. Darker shading represents higher values.

Figure 7: Cutoff value vs. fraction above cutoff curve for the block average images in Figure 6. The solid lines are the training images, dashed lines the simulations. Black is 30% high values, red 35%, green 50%.
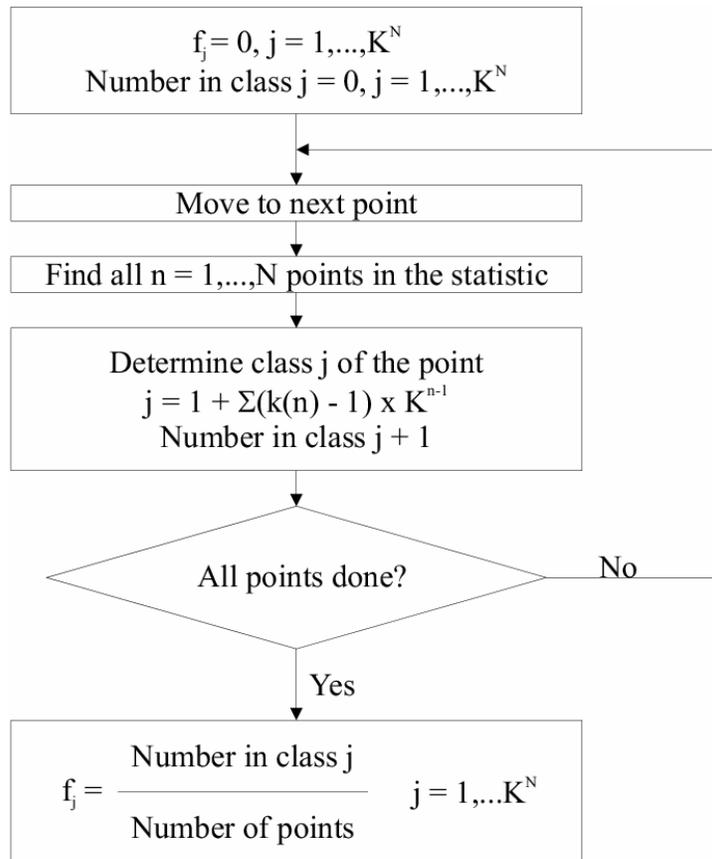
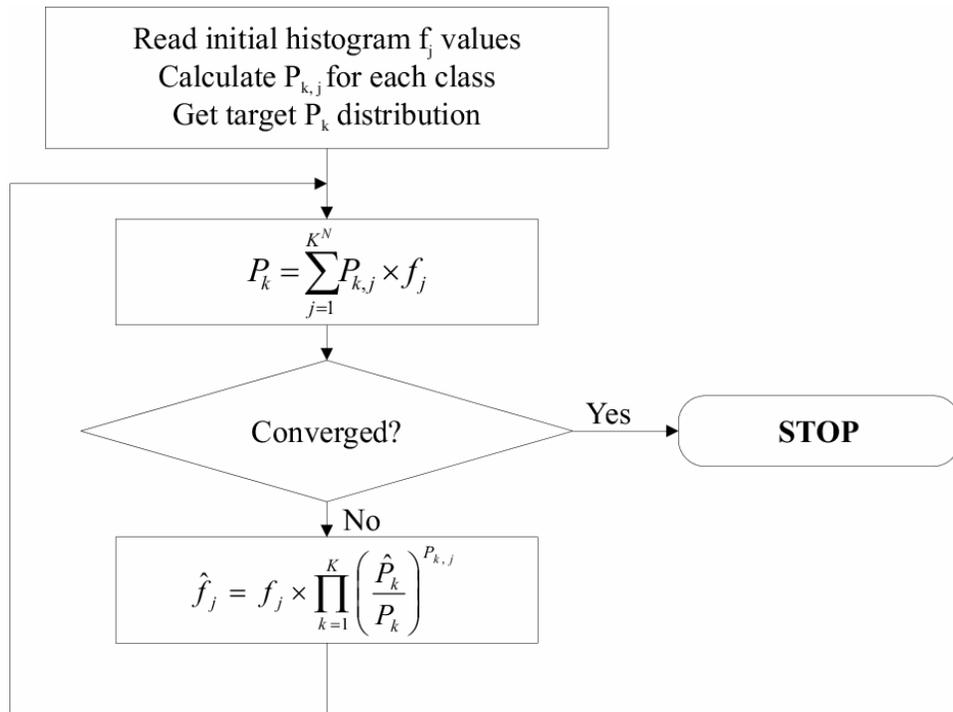**Figure 8**: The procedure used to calculate multiple-point histograms for the example.

Figure 9: The scaling procedure used. It involves iterating Equation (5) until the resulting P_k values converge with those desired.