Comparing Effects of Aggregation Methods on Statistical and Spatial Properties of Simulated Spatial Data

Ling Bian and Rachael Butler

Abstract

Spatial data aggregation is widely practiced for "scaling-up" environmental analyses and modeling from local to regional or global scales. Despite acknowledgments of the general effects of aggregation, there is a lack of systematic comparison between aggregation methods. The study evaluated three methods – averaging, central-pixel resampling, and median – using simulated images. Both the averaging and median methods can retain the mean and median values, respectively, but alter significantly the standard deviation. The central-pixel method alters both statistics. The statistical changes can be modified by the presence of spatial autocorrelation for all three methods. Spatially, the averaging method can reveal underlying spatial patterns at scales within the spatial autocorrelation ranges. The median method produces almost identical results because of the similarities between the averaged and median values of the simulated data. To a limited extent, the central-pixel method retains contrast and spatial patterns of the original images. At scales coarser than the autocorrelation range, the averaged and median images become homogeneous and do not differ significantly between these scales. The central-pixel method can induce severe spatially biased errors at coarse scales. Understanding these trends can help select appropriate aggregation methods and aggregation levels for particular applications.

Introduction

Spatial data aggregation is widely used in environmental analyses and modeling such as in ecology and hydrology that have moved from local scaled modeling to larger regions (Ebleringer and Field, 1993; Gupta *et al.*, 1986). During an aggregation process, the original spatial data are reduced to a smaller number of data units (points, lines, polygons, or pixels) for the same spatial extent. As a result, each aggregated data unit represents a larger area than the original units. The aggregated data are often referred to having a coarser spatial resolution. In an era that emphasizes global scaled research, data aggregation is widely practiced primarily for "scaling up" environmental analyses or models from local to regional or global scales. Spatial data available at finer resolutions need to be aggregated to represent the spatial characteristics (spatial pattern, spatial autocorrelation, etc.) at corresponding scales.

The aggregation process may alter the statistical and spatial characteristics of the data. When aggregated data are used as input to analyses or models, the output of these analyses or models may be affected, i.e., outputs differ when input data of different resolutions are used. This effect is no longer an unfamiliar phenomenon to the GIS, remote sensing, and other science communities that use spatial data (Quattrochi and Goodchild, 1997). Despite the fact that general effects of aggregation are often acknowledged, there is a lack of systematic evaluation of the effects caused by different aggregation methods. Studies that require aggregation often employ the most convenient method without taking potential effects into account. As a result, this may jeopardize the integrity of the studies as well as of any subsequent decision-making processes.

For scientific inquiry, aggregating data to a coarser resolution is often preferred, because certain spatial patterns will not be revealed until the data are presented at a coarser scale (Seyfried and Wilcox, 1995; Zhang and Montgomery, 1994; De Cola, 1994; Bian and Walsh, 1993; Brown *et al.*, 1993; Vieux, 1993; Levin, 1993; Lam and Quattrochi, 1992; Stoms, 1992; Turner *et al.*, 1989; Nellis and Briggs, 1989; Mark and Aronson, 1984). In other circumstances, it is feared that data aggregation could cause information loss, thus having a negative impact on a study. All aggregation methods lose details, but some methods can retain statistical characteristics of the original data better than others. Similarly, some methods may help reveal new spatial patterns better than they can maintain statistical characteristics of the original data. This notion adds another dimension to a systematic evaluation of aggregation effects.

There is a growing literature reporting on the effects of data aggregation under the general topic of scale effects (Wolock and Price, 1994; Zhang and Montgomery, 1994; De Cola, 1994; Bian and Walsh, 1993; Brown *et al.*, 1993; Vieux, 1993; Lam and Quattrochi, 1992; Stoms, 1992; Turner *et al.*, 1989; Nellis and Briggs, 1989; Mark and Aronson, 1984). Most of this type of work appears in the literature of GIS, remote sensing, environmental modeling, or other sciences. The modeling communities tend to focus their attentions on aggregation effects on model output. This narrow focus fails to separate inherent errors from operational errors. Inherent errors are those that are carried in the data before the data are entered into analyses or models, while operational errors occur during analyses or modeling (Walsh, 1989; Burrough, 1986). If aggregated data are used as input for analyses and modeling, errors induced during aggregation are inherent errors. A successful identification of inherent errors can help devise means to compensate for those errors. For example, knowing that aggregated DEM data caused the mean of a wetness index to increase, Lammers *et al.* (1997) subtracted a constant from the

Department of Geography, State University of New York, Buffalo, NY 14261-0023.

mean, leading to a much improved hydrological modeling result.

The majority of studies that have paid attention to identifying inherent errors (caused by aggregation) use empirical data. The knowledge learned is often on an ad hoc basis. For a simple image such as one with row crop on rolling hills, it is possible to link scales of the two features to changes in statistics at a particular aggregation level (e.g., transition points in a semivariogram). For most empirical images, spatial patterns are often too complex to pin down exact spatial features that might correspond to particular statistical changes. In dealing with this problem, simulated data can be a better option for systematic evaluation of aggregation effects in order to ultimately understand the "real world" patterns. Some recent works showed advantages of simulation in providing better control over statistical and spatial characteristics of data (Hunter and Goodchild, 1997; Arbia *et al.*, 1996; Heuvelink, 1992). In particular, Arbia *et al.* (1996) studied data aggregation effects using simulated images. However, their simulated images were restricted to only two values, 0 and 1. All the aforementioned simulations implemented spatial autocorrelation but with only one "structure," which implies that an image is dominated by only one type of spatial feature. For example, an image with row crop on rolling hills would involve two structures, the row crop with a shorter range of spatial autocorrelation and the rolling hills with a longer range. It is necessary to move one step forward to create simulations that are closer to a natural image, i.e., images that have a greater number ($>> 2$) of data values and multiple structures of spatial autocorrelation.

Objectives

The objective of this research is to compare and evaluate three widely used spatial data aggregation methods using simulated data with multiple structures of spatial autocorrelation. The three aggregation methods—averaging, central-pixel resampling, and median—are commonly used in spatial analysis in many scientific disciplines, and are used for aggregating both GIS and remotely sensed data. The simulated images contain a great range of data values and two structures of spatial autocorrelation so that they more closely resemble natural images of biophysical phenomena. The aggregation effects were evaluated by changes in statistical and spatial distribution, as well as by the changes of errors across aggregation levels. Special attention is given to changes at levels that correspond to the ranges of spatial autocorrelation embedded in the data.

The remainder of the paper presents the following sections of discussion. The first one details the data simulation design and implementation. The second section describes aggregation procedures employed on the data. The third section summarizes and discusses the aggregation results, and the fourth section discusses implications of the results to scaling-up analyses and modeling. A final section concludes the findings.

The Simulation Design

The basic spatial unit used for the simulation is typical of raster data, the pixel. The importance of raster data is deeply rooted in the success of remote sensing (Couclelis, 1992; Peuquet, 1988) and the DEM. These two types of spatial data, along with many other types of raster data, are widely used in geography as well as in many areas of natural science. The raster data are ideal for studying data aggregation because of their regular and uniform geometry (e.g., pixels share identical size and shape). With raster data, spatial resolution can be readily measured and presented. In addition, the effects of data aggregation can be easily observed by varying the data resolution and comparing the resultant sta-

tistical and spatial characteristics. The raster data are well suited to the intended study.

There are many models for simulating spatial phenomena. Burrough (1986), following the regionalized variable theory, argued that the spatial variation of any variable can be expressed as the sum of three major components:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon'' \qquad (1)$$

where $Z(x)$ is the value of variable Z at location x. The first component, $m(x)$, is a general trend (a tilted plane, for example) that can be described by a linear model. The second component, $\varepsilon'(x)$, is random but it is locally varying and spatially autocorrelated. This component can be described by spatial autocorrelation models. The third component, ε'', is a spatially independent error term, that can be modeled by a random distribution such as Gaussian distribution. This conceptual framework is adopted for simulation in the present study and the general trend is excluded to assure stationarity across the simulated area. Arbia *et al.* (1996) also simulated the same two components but with a spatially autocorrelated error term.

The simulation is implemented using unconditional simulation described in Deutsch and Journel (1998). A simulation process produces a series of realizations of spatial data according to pre-defined parameters for both statistical distribution and spatial autocorrelation. A conditional simulation generates data using parameters of actual sample data. An unconditional simulation creates data without actual sample data and the simulation uses default parameters (e.g., mean = 0). Each realization is the sum of the estimated data component and the corresponding error term. More elaborated treatment of spatial simulation can be found in Deutsch and Journel (1998).

The present study assumes a Gaussian distribution for the data component based on the model detailed by Burrough (1986). Each estimated data value is a possible outcome of the Gaussian random function. The estimated data follow pre-defined spatial autocorrelation. In the present study, the spatial autocorrelation contains two structures, both following a standard spherical model. The expression of the model follows the form defined in Deutsch and Journel (1998): i.e.,

$$\gamma(h) = c \cdot \mathrm{Sph}\,(h/a) = c \cdot [1.5(h/a) - 0.5(h/a)^3]\ \text{if}\ h \leq a \qquad (2)$$
$$= c \qquad\qquad\qquad\qquad\quad \text{if}\ h \geq a$$

where semivariance $\gamma(h)$ is a function of the spatial lag h, a is the range of spatial autocorrelation, and c is the sill. Once the spatial lag reaches the range, the semivariance $\gamma(h)$ becomes a constant. Among several spatial autocorrelation models, the spherical model has the shape for which the range of spatial autocorrelation can be most easily identified in a semivariogram by visual examination. This is ideal for the intended study. The nugget effect is set to zero. The first structure of the spatial autocorrelation has a spatial autocorrelation range of 30 pixels and the second structure has a range of 10 pixels. The error term is simulated such that it is independent of the estimated data component and it also follows a Gaussian random function.

The total simulated area is 512 by 512 pixels, and the simulation generates a total of 30 realizations to assure an adequate number of observations for summary statistics. Table 1 lists the basic statistics of the simulated images. The simulation is implemented using the geostatistical software package GSLIB (Deutsch and Journel, 1998). The choice of the area size is primarily constrained by technical difficulties. The unconditional simulation algorithm (kriging) is computationally intensive. The present study did attempt to simulate 1024- by 1024-pixel areas, but an experimental run indicated that simu-

TABLE 1. SUMMARY STATISTICS OF THE 30 SIMULATED IMAGES. THE MEAN, STANDARD DEVIATION, MAXIMUM, AND MINIMUM VALUE FOR EACH OF THE 30 IMAGES WERE COMPUTED (ROWS). THE SUMMARY STATISTICS PRESENTED AS THE MAXIMUM, MINIMUM, AND MEAN OF THE FOUR SETS OF STATISTICS OVER THE 30 IMAGES ARE PRESENTED IN COLUMNS.

	Maximum	Minimum	Mean
Means (30)	0.1233	−0.0606	0.0124
Standard Deviations (30)	1.3900	1.3201	1.3471
Maximum Values (30)	7.1950	5.1956	5.9649
Minimum Values (30)	−5.1837	−6.8572	−5.9399

lating such sizes demanded super-computing power. The area of 512 by 512 pixels is the largest area possible within the capacity of Unix workstations. A 512- by 512-pixel area is adequate for the intended study. It contains a number of 30-pixel structures across the area so that the image reasonably resembles a natural image, while the number of structures is within a manageable range.

Aggregation Procedures

Three aggregation methods — averaging, central-pixel resampling, and median — are used for aggregating the simulated data images. All the methods extract a value over an n by n window in the original image. The windows are adjacent but do not overlap (Figure 1). The averaging method uses the average value over the n by n window. This method may be deemed more appropriate for aggregating remotely sensed images, because a pixel value is assumed to be the integrated value over the corresponding area on the ground. The central-pixel resampling method takes the original value of the central pixel of the n by n window. This method is more commonly used in modeling communities such as hydrology (Wolock and Price, 1994). The median method takes the median of the n by n window. This method is perhaps less used in practical works but it is easy to implement and less sensitive to extreme data values. It is examined in the present study for its potential as an alternative for the averaging method. Initially, a mode method was attempted which takes the mode value of a window. For image-like data, however, either all data values in a window are unique so no mode value can be identified, or a handful of pixels share identical values to yield a mode value, but this situation does not carry the meaning of mode. A C++ code was developed in-house to implement the three aggregation methods.

The three methods are applied to the same 30 simulated images to compare differences among the methods. The aggregation operates at ten levels, using 3- by 3-, 9- by 9-, 11- by 11-, 21- by 21-, 31- by 31-, 41- by 41-, 51- by 51-, 61- by 61-, 71- by 71-, and 81- by 81-pixel window sizes. All windows are positioned to begin from the upper left-most pixel of the 512- by 512-pixel area. At each level, data are aggregated directly from the original images (e.g., from 1 by 1 to 51 by 51) instead of from a previous aggregation (e.g., from 41 by 41 to 51 by 51). Window sizes 9 by 9, 11 by 11, and 31 by 31 are included to examine the behavior of data at aggregation levels corresponding to the ranges of spatial autocorrelation. Window sizes around the autocorrelation ranges, 21 by 21 and 41 by 41 pixels, are also included in case the autocorrelation ranges deviate from the simulation design due to the errors embedded in the simulation. The small window size (3 by 3) and the large window sizes (51 by 51 to 81 by 81) provide observations to monitor the trend of change at aggregation levels finer and coarser, respectively, than the autocorrelation ranges. With aggregation, the total number of pixels decreases and each pixel represents a larger
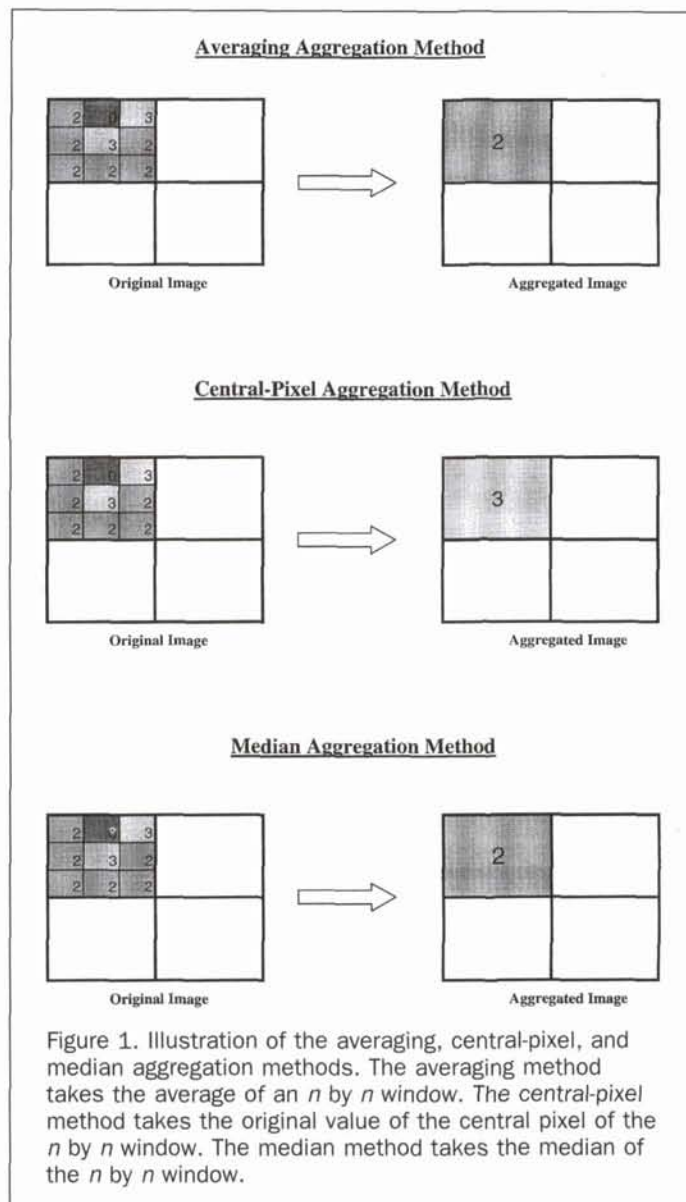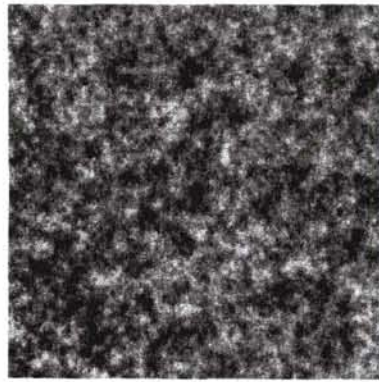

Figure 1. Illustration of the averaging, central-pixel, and median aggregation methods. The averaging method takes the average of an *n* by *n* window. The central-pixel method takes the original value of the central pixel of the *n* by *n* window. The median method takes the median of the *n* by *n* window.

area (Figure 2). The 512 by 512 pixels at the original data level are reduced to 6 by 6 pixels at the last level of the 81- by 81 pixel window size. Further aggregation leaves too few pixels to support statistical analysis.

The statistical distribution and spatial distribution are evaluated and compared between the three methods across the ten aggregation levels. The distributions are evaluated by three sets of treatment: (1) the mean and standard deviation of the aggregated data plotted against window size for all methods, (2) histograms of all levels for all methods to support the statistics in (1), and (3) images of aggregated data by all methods at each aggregation level to evaluate the changes in spatial pattern during aggregation.
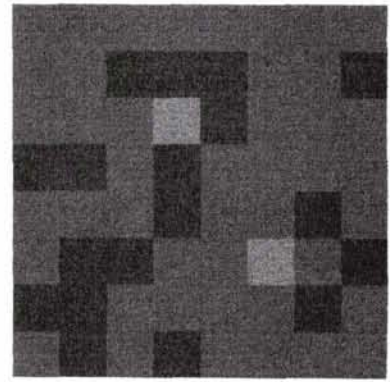
The spatial details lost during aggregation are considered to be aggregation errors. Error images are created by subtracting the aggregated values from the originally simulated values in each window. The resultant error images maintain the original size of 512 by 512 pixels. The statistical and spatial characteristics of errors are evaluated by three sets of treatments: (1) mean and standard deviation of errors to evaluate statistical changes of errors across all aggregation levels for all meth-
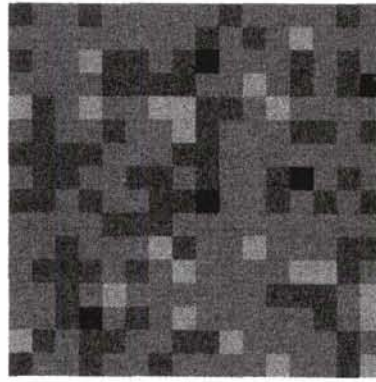
(a) Original   (b) 31x31 pixels   (c) 61x61 pixels

Figure 2. Illustration of aggregated images. Presented are (a) an original image, (b) an aggregated image using a 31- by 31-pixel window by the averaging method, and (c) an aggregated image using a 61- by 61-pixel window by the averaging method.
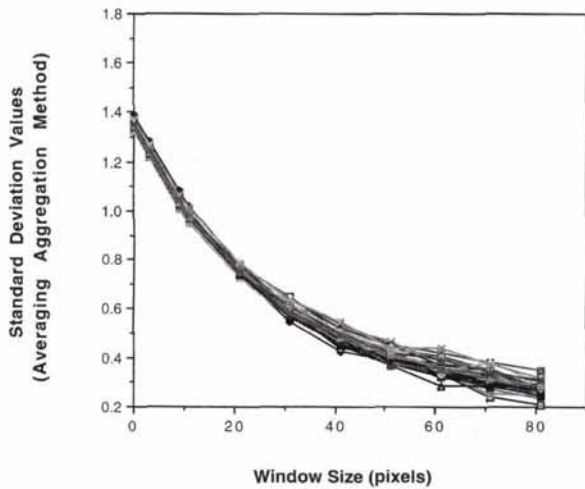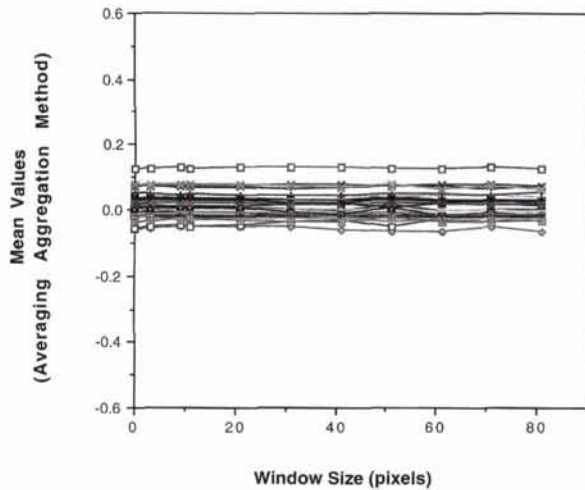

Figure 3. Means and standard deviations of the averaging images for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.
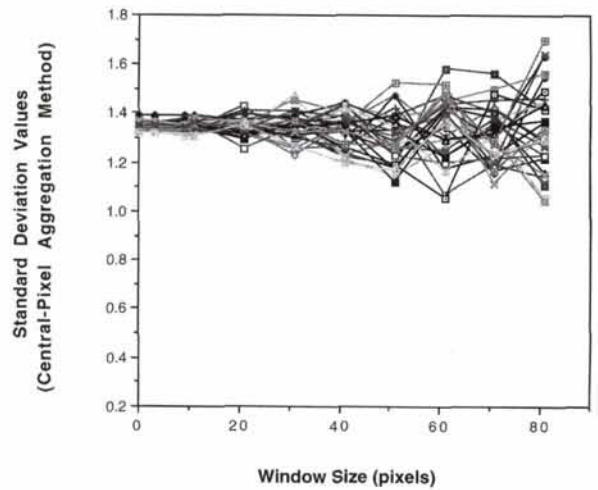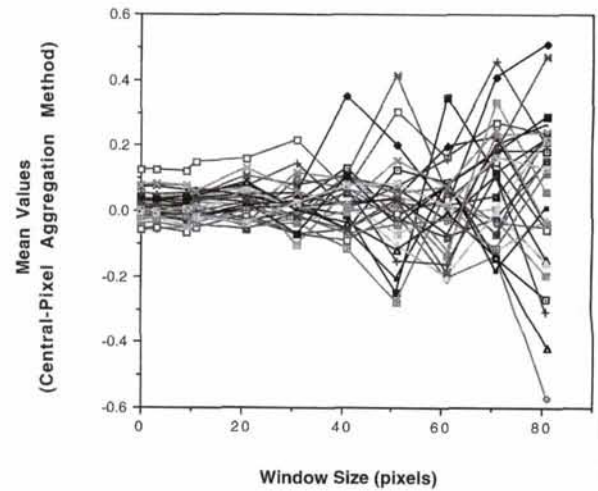

Figure 4. Means and standard deviations of the central-pixel images for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.
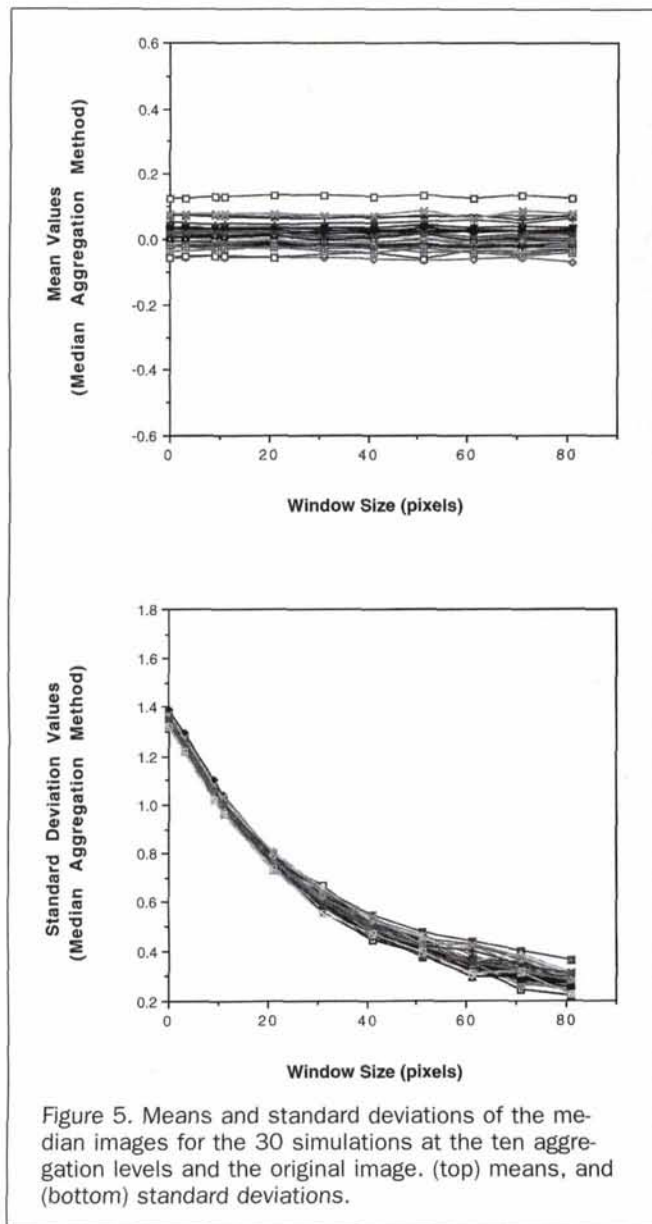
Figure 5. Means and standard deviations of the median images for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.

central-pixel images (Figure 4) show a rather chaotic pattern. Up to the 10-pixel window size (9 by 9, 11 by 11), the means remain almost the same as those of the original images. At window sizes greater than 30 pixels, the values spread to a wide range.

The differences in aggregation mechanisms among the three methods cause the fundamental difference in statistics of the aggregated data. Both the averaging and median methods aggregate based on data values, and the values are confined to the mid-range. The central-pixel resampling is based on location, which changes with window size. The aggregation results are a systematically sampled subset of the original data, and their values are expected to be less confined. This may explain the spread of the means for the central-pixel method (Figure 4). The spatial autocorrelation further modifies the general statistical trend. Within windows smaller than the spatial autocorrelation ranges, spatial autocorrelation is present. The means are similar to the original neighboring values within the window. With windows greater than the spatial autocorrelation range, the statistical distribution approaches the global distribution. An aggregated image based on central pixels becomes a mere subset of the original image. Its mean value varies with the location of central pixels as the window size changes.

The standard deviations show a similar division among the three methods. The averaging and the median methods present a decreasing trend with increasing aggregation level (Figures 3 and 5, respectively). This is expected because both

ods, (2) images of errors at all aggregation levels for all methods to examine the changes in spatial characteristics of the errors, and (3) semivariograms of errors to support illustrations in (2). A total of 2160 images are analyzed, including 11 aggregation levels (including the original images), three methods, 30 realizations, and a data and an error image for each level, method, and realization. It takes more than three gigabytes of space to store the 2160 data and error images!

Aggregation Results and Analysis

The means and standard deviations (Figures 3, 4, and 5) summarize the statistical changes through aggregation. The means of the averaging and median images remain essentially constant across all aggregation levels (Figures 3 and 5, respectively). Theoretically, the means of the averaging images should be absolutely constant. The mild variations shown in Figure 3 are due to an "edge effect," which is produced by the inability of the window size to divide evenly into 512 pixels. The pixels that do not make up a whole window (edge pixels) are left out of the computation, thus slightly shifting the statistics of an image. In a sharp contrast, the means of the
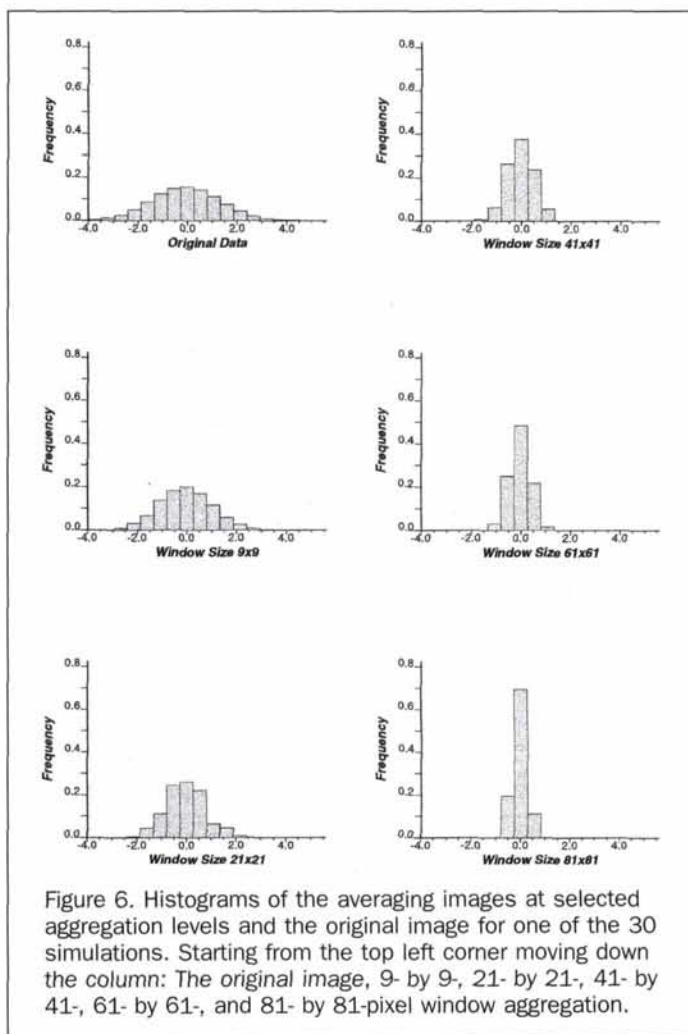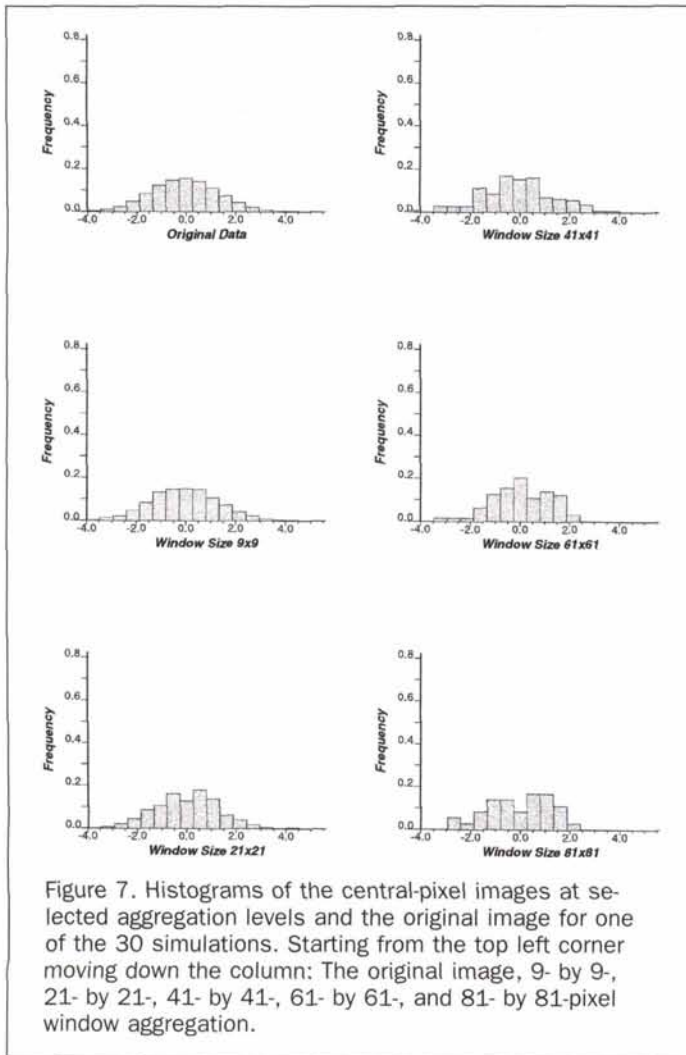

Figure 6. Histograms of the averaging images at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.

Figure 7. Histograms of the central-pixel images at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.

tral-pixel method displayed in Figure 7 shows similarities between the original and the 9- by 9-pixel window as well as the fluctuation after the 41- by 41-pixel window.

Changes in spatial patterns are presented in Figures 9, 10, and 11 for aggregated images produced by the three methods. The averaging method, functioning as a low-pass filter, reveals the underlying spatial patterns and loses contrast (Figure 9). At window sizes greater than 30 pixels, the averaged values quickly approach the global mean value, and these values appear similar between large window sizes. Because of the similarity between the averaged and median values, the median images appear almost identical to the averaging images (Figure 11). Figure 10 shows the central-pixel images. By visual examination, the contrast as well as the basic patterns (high and low values and their relative locations) of the original images seem to be better kept than the other two methods. The patterns become increasingly random at window sizes larger than 30 pixels.

The spatial details lost during aggregation represent aggregation errors. Figure 12 presents statistical summaries of errors produced by the averaging method. The means of errors (Figure 12) vary about zero and the range of the variation slightly increases with window size. The errors are small because they are the differences between the original values and their local averages. Standard deviations of the errors level off around the 30-pixel window size. At this window size and larger, the averaged values are similar to the global mean. As a result, the standard deviations of the errors are similar among these windows and approach the global standard deviation (Table 1).

Figure 8. Histograms of the median images at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.
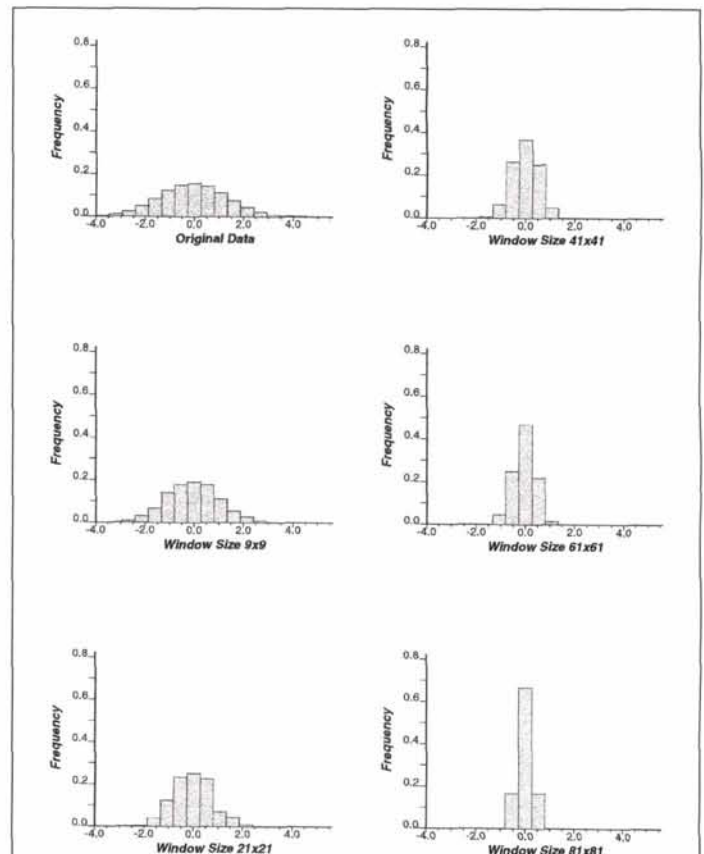
methods eliminate the low frequency values at both ends of a histogram (the large and small values), resulting in a smaller standard deviation. This effect is more obvious at larger window sizes, where a greater range of values are likely to be averaged or eliminated, producing successively "taller and tighter" histograms (Figures 6 and 8) and lower standard deviations. It is important to note the following observations. First, the decrease in standard deviation is non-linear. This could be affected by spatial autocorrelation. The decreasing rate remains almost constant up to ten-pixel window sizes and slows down around the 30-pixel window. It is possible that, when the statistics of large windows (>30 pixels) are approaching the global statistics, the standard deviations do not differ greatly with increasing window size. The non-linear trend was also reported by Arbia et al. (1996) but the study did not provide any explanation.
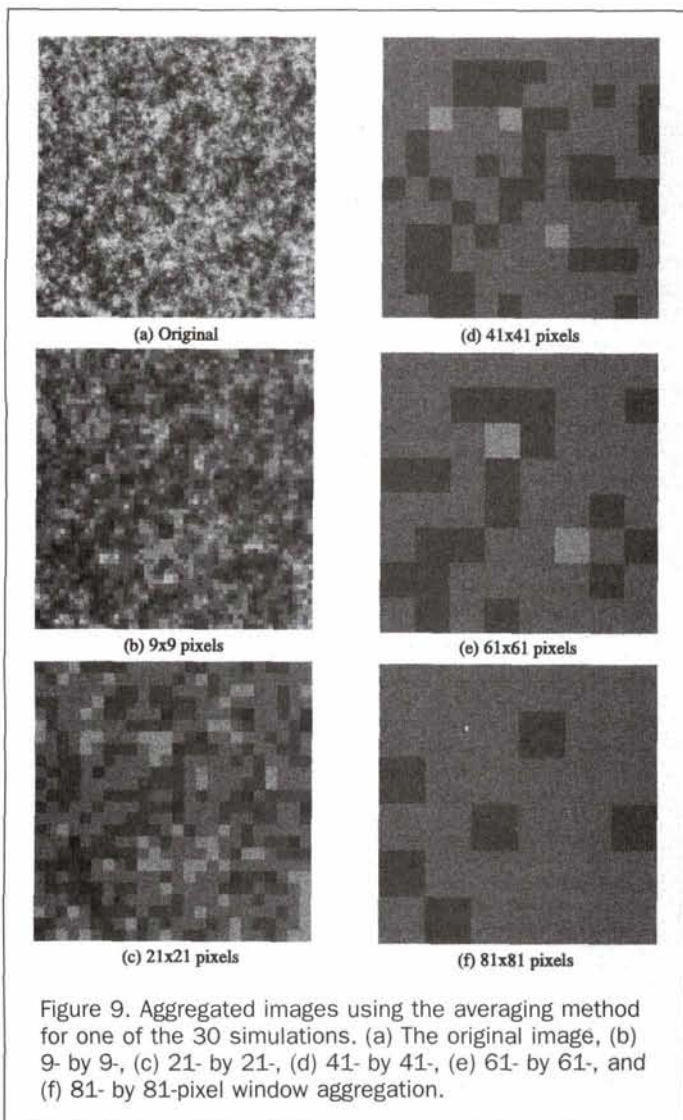
The central-pixel method presents a different trend (Figure 4). The standard deviations fluctuate about standard deviations of the original images (Table 1) across aggregation levels, and the range of variation increases steadily with window size. The fluctuation stems from the varying subsets of data with changing window size. In addition, the small number of pixels can also enhance the fluctuation at large window sizes. Up to window sizes of ten pixels, the standard deviations remain similar to those of the original images. Up to the window size of 40 pixels, the values are confined to a relatively narrow range. Both of these observations can be attributed to the presence of spatial autocorrelation. Histograms of the cen-

Figure 9. Aggregated images using the averaging method for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.

Figures 13 and 14 present spatial pattern changes of errors. Error images (Figure 13) of small window sizes, reflecting local errors, are noisy and vary at fine scales. A global pattern emerges when window size becomes greater than the spatial autocorrelation ranges. The global pattern appears to be a mirror image of the original image and the pattern remains similarly for larger window sizes. In order to examine changes in spatial autocorrelation, semivariograms in Figure 14 summarize the spatial behavior of the errors through aggregation. The range of spatial autocorrelation of the errors gradually approaches 30 pixels where the errors change from a local to a global pattern. Accordingly, semivariance values increase quickly at small window sizes and eventually reach the value of the global variance. This corresponds to the behavior of the standard deviation illustrated in Figure 12.

Figures 15, 16, and 17 present the statistical summary, spatial pattern, and semivariograms of errors, respectively, for the central-pixel aggregation method. The means (Figure 15) have a much greater value range, and the standard deviations have both higher values and a greater value range than those of averaging errors (see Figure 12). Because the central-pixel errors depend on locations of the central pixel which vary with window size, the greater value range of the means and standard deviations is expected. Figure 16 shows that the central-pixel error images appear more "blocky" than the averag-
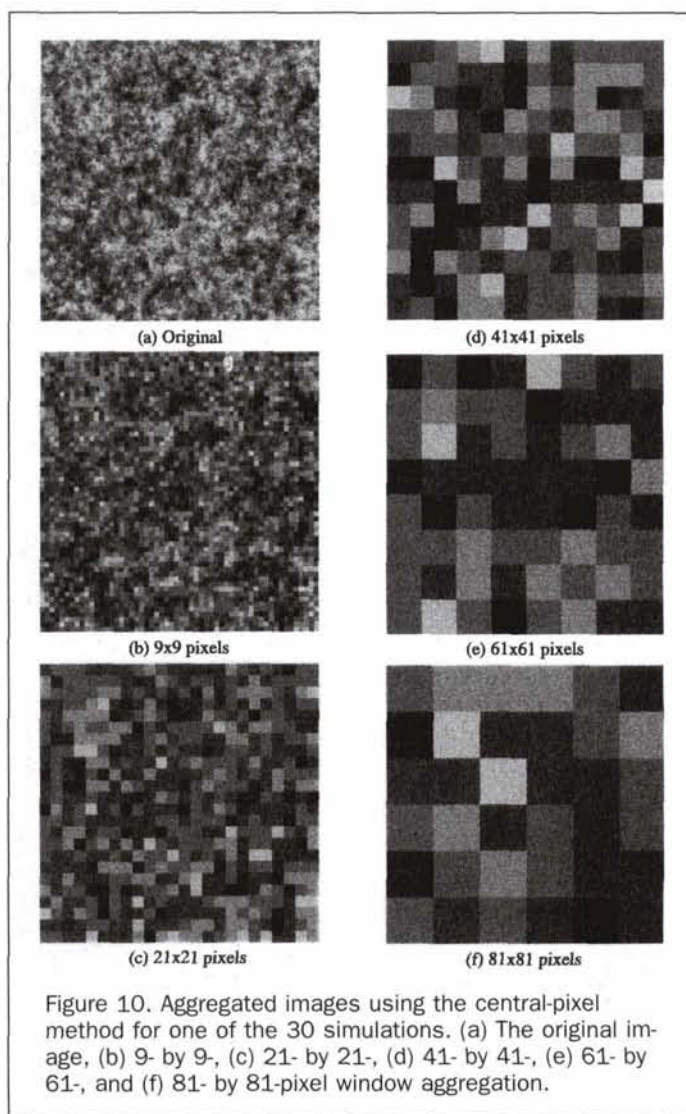
ing error images (Figure 13) when varying values are subtracted from the original image. Semivariograms in Figure 17 show both similarities and differences from those of averaging errors (Figure 14). The range of spatial autocorrelation reaches that of the original image at the 30-pixel window. The semivariograms deviate significantly from that of the original image at larger window sizes, corresponding to the spatial pattern presented in Figure 16.

The statistical and spatial characteristics of median errors are presented in Figures 18, 19, and 20. The statistical characteristics represented by means and standard deviations (Figure 18) and spatial characteristics represented by error images (Figure 19) and semivariograms (Figure 20) are almost identical to those of the averaging errors (Figures 12, 13, and 14). This resemblance stems from similarities between averaged and median values in a window for the data simulated in the present study.

Discussion

Summary of the Observations
The summary focuses on the ability of each method in preserving the statistical and spatial characteristics of the original images, as well as the characteristics of aggregation errors. The averaging and median aggregation methods can maintain the

Figure 10. Aggregated images using the central-pixel method for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.
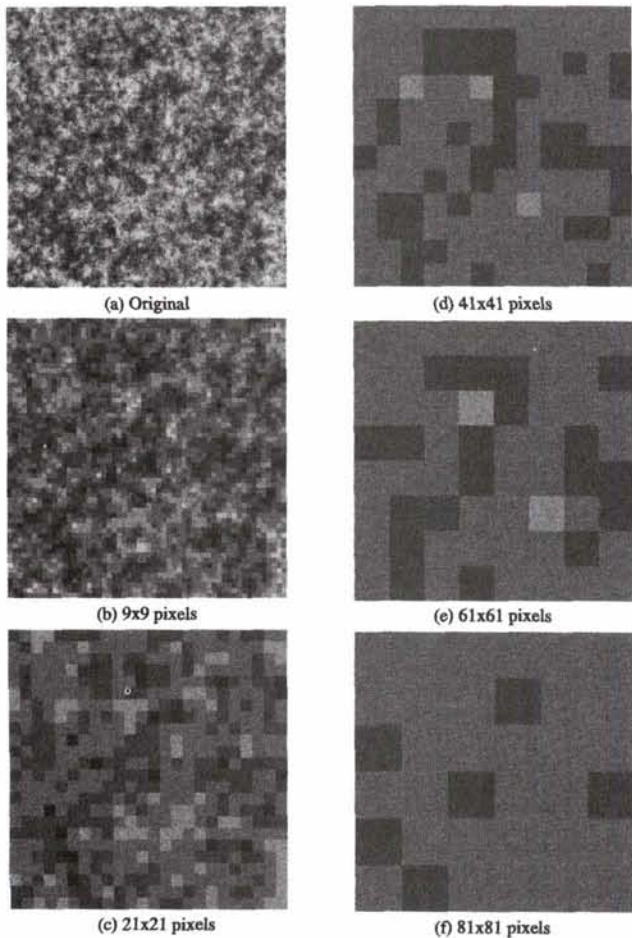
Figure 11. Aggregated images using the median method for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.

Implications of Results

The averaging method can preserve the mean values of original images across all levels of aggregation. This is important particularly for physical process models, many of which require averaging of a variable over an area and then use the mean values as input for modeling. For example, average values of land surface roughness over sub-basins are often used as an input parameter in watershed hydrologic models. The potential cost of the averaging method is the reduced standard deviation. If a model is sensitive to the range of variable values, using an averaging method may induce significant errors. The magnitude and range of errors are affected by the spatial autocorrelation present in the data. Averaging within spatial autocorrelation ranges can confine aggregation errors to a low magnitude and random spatial patterns. Knowing how the decreasing rate of the standard deviation may change, it might be possible to devise a means to "restore" the standard deviation for those models that are sensitive to variable value ranges. Effective means to achieve the goal are yet to be investigated.

Spatially, the averaging method reveals underlying spatial

mean or median of the original images across aggregation levels and also produce decreasing standard deviations. The rate of decrease seems to be associated with the spatial autocorrelation. These observations are shown in Figures 3 and 5. Spatially, both methods lead to increasingly homogeneous images as window size becomes coarser than the spatial autocorrelation range, indicated by the patterns shown in Figures 9 and 11. The standard deviations of both averaging and median errors increase up to the window sizes corresponding to the spatial autocorrelation, as displayed in Figures 12 and 18. Spatially, the error images of both methods change from a noisy, local pattern to a global pattern at the window size of the spatial autocorrelation, and this observation is supported by patterns shown in Figures 13 and 19.

The means and standard deviations of the central-pixel method are less predictable when the location of the central-pixel changes with window size. The presence of spatial autocorrelation seems to help confine the variations (Figure 4). To a limited extent, the method can better maintain data values and basic spatial patterns of the original images within the spatial autocorrelation range. The central-pixel errors have greater magnitude and variation than the other two methods (Figure 15). Spatially, the errors are imprinted heavily by the random patterns involved in the central-pixel images (Figures 10 and 16), especially at window sizes greater than the spatial autocorrelation.
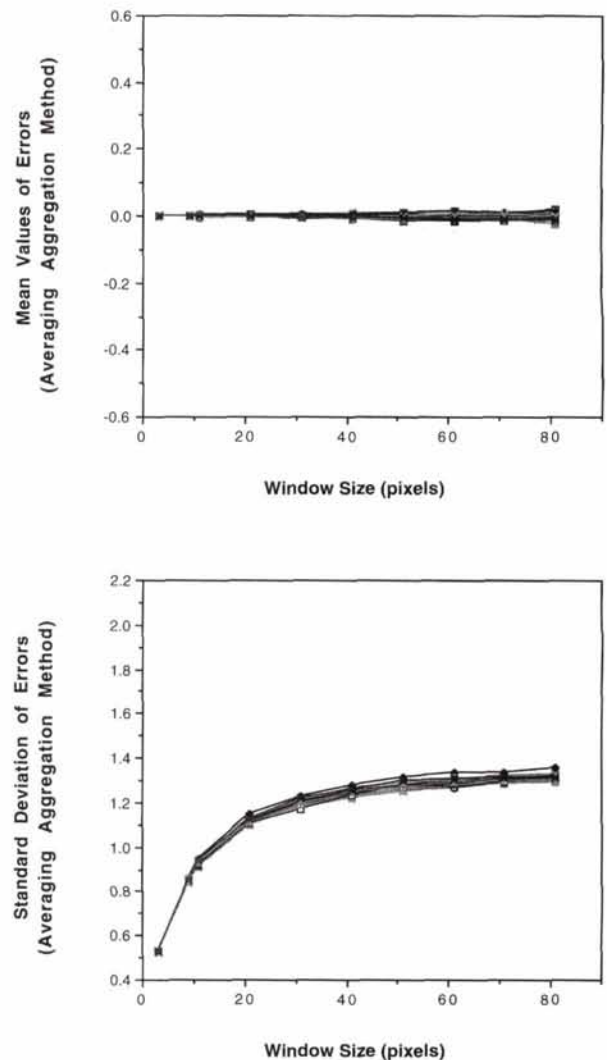

Figure 12. Means and standard deviations of the averaging errors for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.
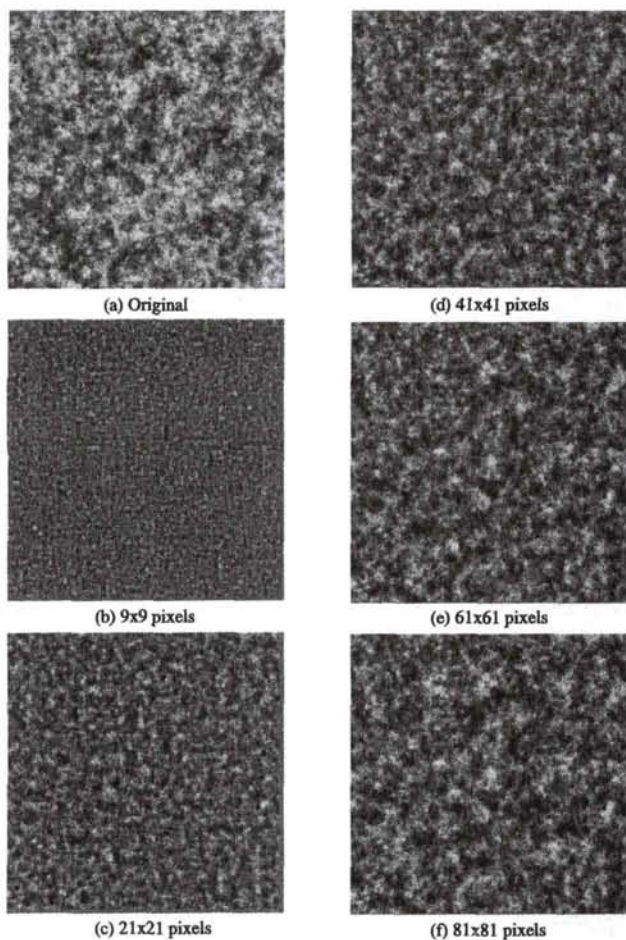
Figure 13. Images of the averaging error at selected aggregation levels and the original image for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.

ing spatial patterns. Resampling central pixels beyond the autocorrelation range may induce unpredictable spatial variation (Figure 10) and spatially biased errors (Figure 16). The latter can be severe. Generating a large number of samples or multiple sample sets may help alleviate the randomness and the subsequent effects.

Data aggregated by the median method share many statistical and spatial similarities with the averaging method. Greater differences are expected for empirical data because the data simulated in the present study are close to an ideal normal distribution so that the average and median are close. Differences can occur in aggregated data and in aggregation errors. Subtracting the median from original values in a window differs from subtracting the average in the window. The former produces greater (if not equal) remainders, or aggregation errors termed in the present study. If the median method does not do better than the averaging method in preserving the statistical and spatial characteristics of the original data, and it only leaves greater aggregation errors, then the median method is not necessarily a better substitute for the averaging method in data aggregation. When there is a choice between the two, averaging is a better option.

Although the ten-pixel spatial autocorrelation is implemented in the simulated images, its effects are not clear in all situations. The effects are more apparent for the central-pixel method in statistics (Figure 4) and semivariograms of errors (Figure 17). It is possible that the smoothing effect of the aver-

patterns (as opposed to original patterns) within the spatial autocorrelation ranges (Figure 9). Averaging beyond the spatial autocorrelation ranges may not reveal any new actual patterns. Instead, the difference between patterns across these aggregation levels may become a statistical artifact (Figure 9). For models that use averaged values and are sensitive to spatial patterns, using data averaged over one resolution coarser than the spatial autocorrelation range may not induce more errors than data averaged over another resolution coarser than the autocorrelation range. This serves as a worst-error scenario if averaging must be done at a coarse resolution.

The central-pixel method, in contrast, yields a greater yet less predictable variation for both mean and standard deviation. For process models that depend on mean values, and/or are sensitive to ranges of variable values, the central-pixel method may induce more inherent errors than the averaging method. An exception is to resample central pixels within the spatial autocorrelation ranges (the autocorrelation range can be identified using semivariogram, using Moran's I index, or visually examining the size of dominant spatial features in an image). This may keep both the mean and standard deviation close to those of the original data. Spatially, the central-pixel method preserves, to a limited extent, the original values and their relative locations at aggregation levels close to the original. This trait can be useful to models that require maintain-
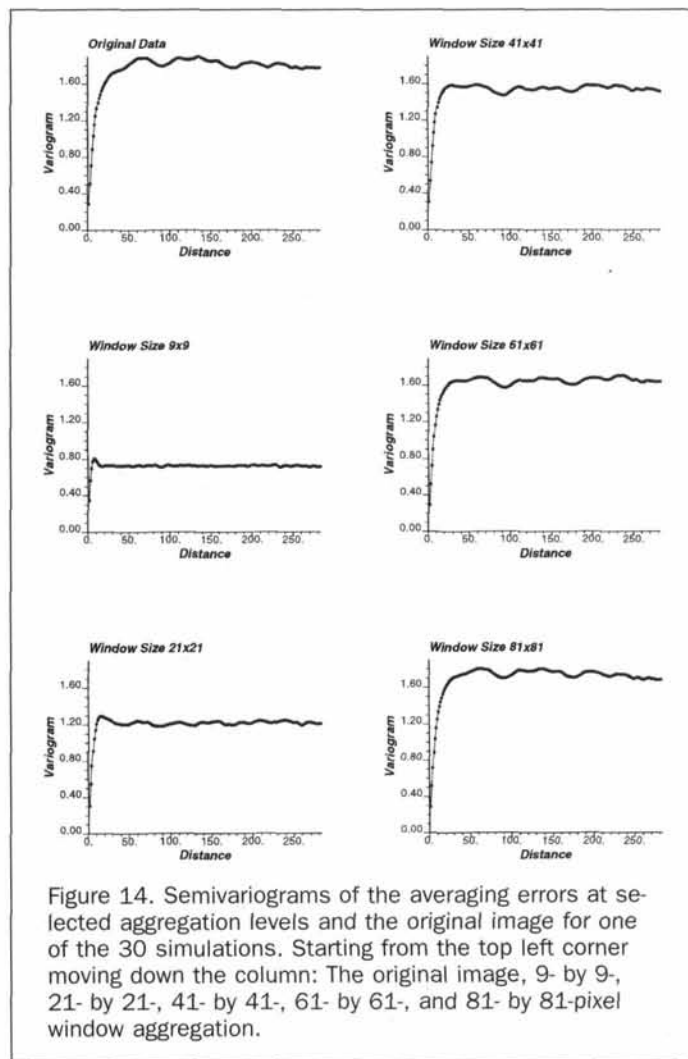

Figure 14. Semivariograms of the averaging errors at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.
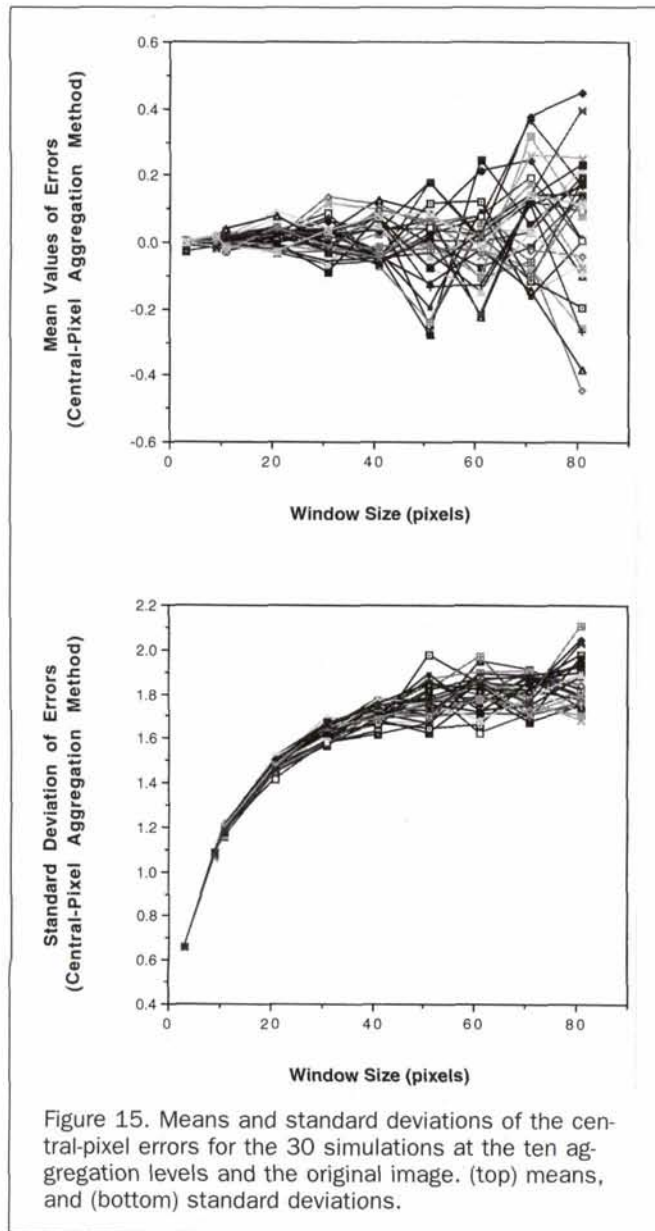
Figure 15. Means and standard deviations of the central-pixel errors for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.

I, etc.) to identify spatial autocorrelation ranges is a necessary step to help choose appropriate aggregation levels. Aggregating within the spatial autocorrelation range can help confine the magnitude or spatial pattern of errors. Knowing the behavior of errors through aggregation can help predict the worst-error scenarios if aggregation must be done at coarse scales. With the error and uncertainty theory, knowledge of error behavior can help access the accuracy of analyses and modeling outcomes (Goodchild and Gopal, 1989). Furthermore, it is possible to devise means to compensate inherent errors. Such examples include simply subtracting a constant from the mean (Lammers *et al.*. 1997) to improve the model performance. Generating multiple subsets of central pixels can help stabilize the variation in the aggregated data induced by the central-pixel resampling method.

Using simulated images provides a better control of statistical and spatial characteristics of the data, and it is suitable for a systematic evaluation of aggregation effects. Results of the evaluation can be compared with a control data set without spatial autocorrelation. Such a comparison may provide more conclusive explanations on effects of spatial autocorrelation in the present study. It will also be useful to examine empirical images that contain apparent, multiple-scaled features to validate results found in the present study.
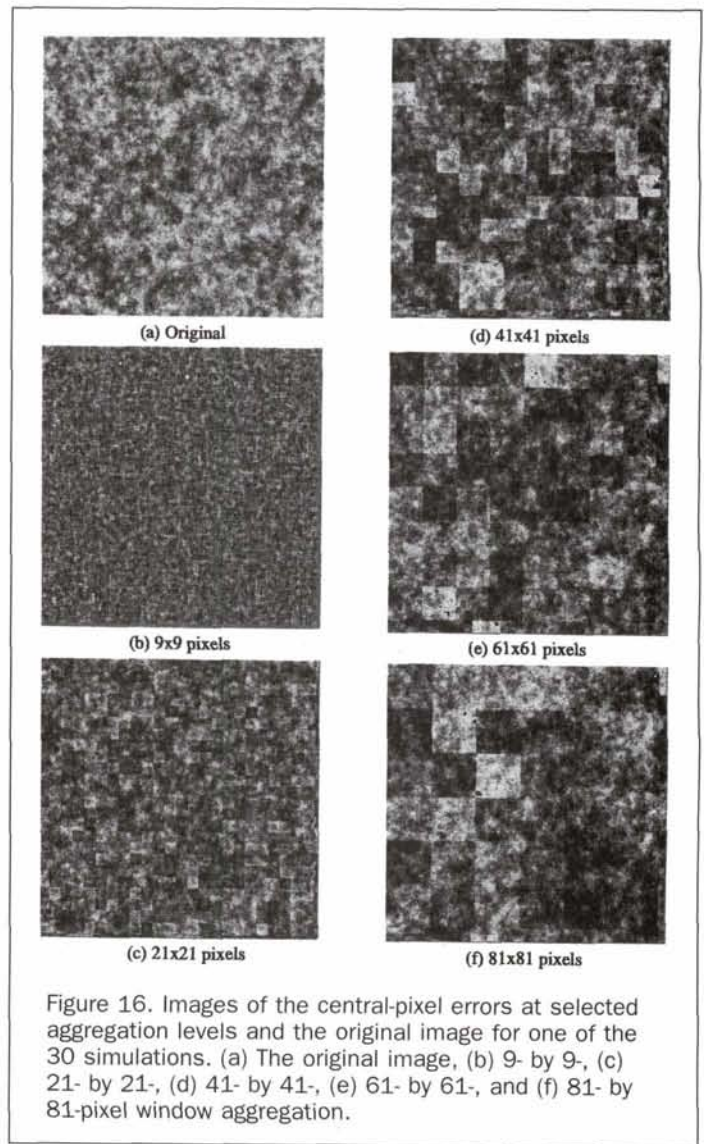

Figure 16. Images of the central-pixel errors at selected aggregation levels and the original image for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.

aging method may have obscured effects of spatial autocorrelation of a shorter range. This observation is consistent with the previous experience of authors using empirical images. Spatial autocorrelation seems to be apparent at 20- or 40-pixel window sizes. It is possible that these two window sizes are transition scales. On the basis of a preliminary experiment on a separate set of simulations, the spatial autocorrelation appeared in the semivariograms tends to extend beyond what is defined in the simulation. The random errors inserted into the simulated data may have caused the shifting.

Conclusions

Among the three methods compared in the present study, the averaging method produces aggregated data and aggregation errors with the most statistically and spatially predictable behavior. The median method produces similar results. The central-pixel method is the least predictable.

Successfully identifying aggregation effects on statistical and spatial properties of spatial data can help manage inherent errors before the data are entered into analyses or models. Using established methods (e.g., semivariograms and Moran's
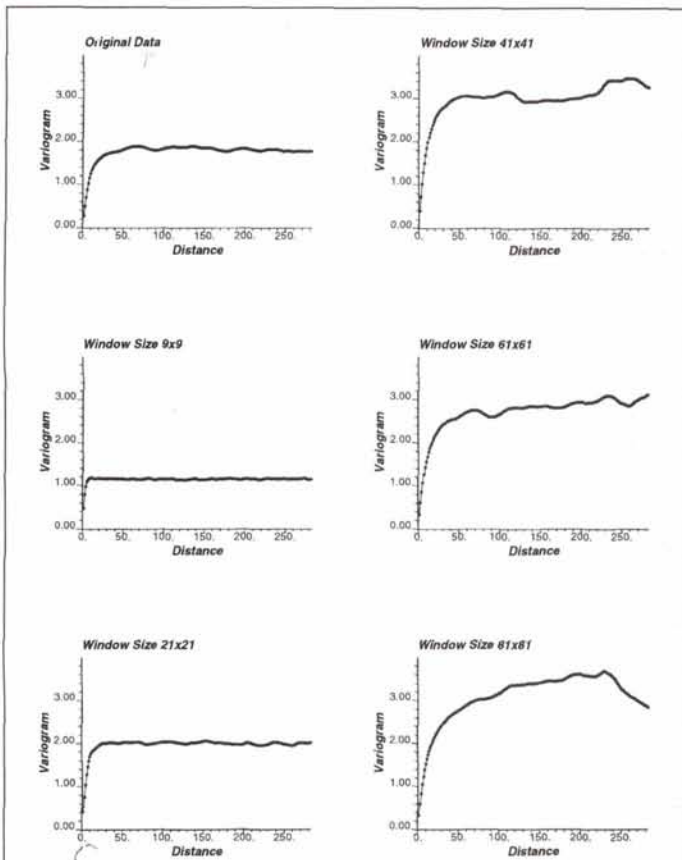
Figure 17. Semivariograms of the central-pixel errors at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.

Another issue that warrants further investigation is the role of multiple structures of spatial autocorrelation in aggregation effects. Although two structures are implemented in the present simulation, only the effect of 30-pixel autocorrelation range is apparent. The impact of the ten-pixel range is not as apparent as expected.

These observations should contribute to a better understanding of aggregation methods and their effects on using multi-resolution spatial data for environmental analyses and modeling. Aggregation effects are an interesting yet complex issue. In-depth understanding of the effects will evolve in response to the challenges of scaling up environmental models for global change research and for a better management of natural resources.

Acknowledgments

References

Arbia, G., R. Benedetti, and G. Espa, 1996. Effects of the MAUP on image classification, *Geographical Systems*, 3(2):123–141.

Bian, L., and S.J. Walsh, 1993. Scale dependencies of vegetation and topography in a mountainous environment of Montana, *Professional Geographer*, 45(1):1–11.

Brown, D.G., L. Bian, and S.J. Walsh, 1993. Response of a distributed watershed erosion model to variations in input data aggregation levels, *Computers and Geosciences*, 19(4):499–509.

Burrough, P.A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment*, Clarendon Press, Oxford, 193 p.

Couclelis, H., 1992. People manipulate objects (but cultivate fields): Beyond the raster-vector debate in GIS, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space* (A.U. Frank, I. Campari, and U. Formentini, editors), Springer-Verlag, Berlin, pp. 65–77.

De Cola, L., 1994. Simulating and mapping spatial complexity using multi-scale techniques, *International Journal of Geographical Information Systems*, 8(5):411–427.

Deutsch, C.V., and A.G. Journel, 1998. *GSLIB-Geostatistical Software Library and User's Guide, Second Edition*, Oxford University Press, New York, 369 p.

Ebleringer, J.R., and C.B. Field, 1993. *Scaling Physiological Processes, Leaf to Globe*, Academic Press, Inc., New York, 388 p.

Goodchild, M.F., and S. Gopal (editors), 1989. *The Accuracy of Spatial Databases*, Taylor & Francis, London, 290 p.
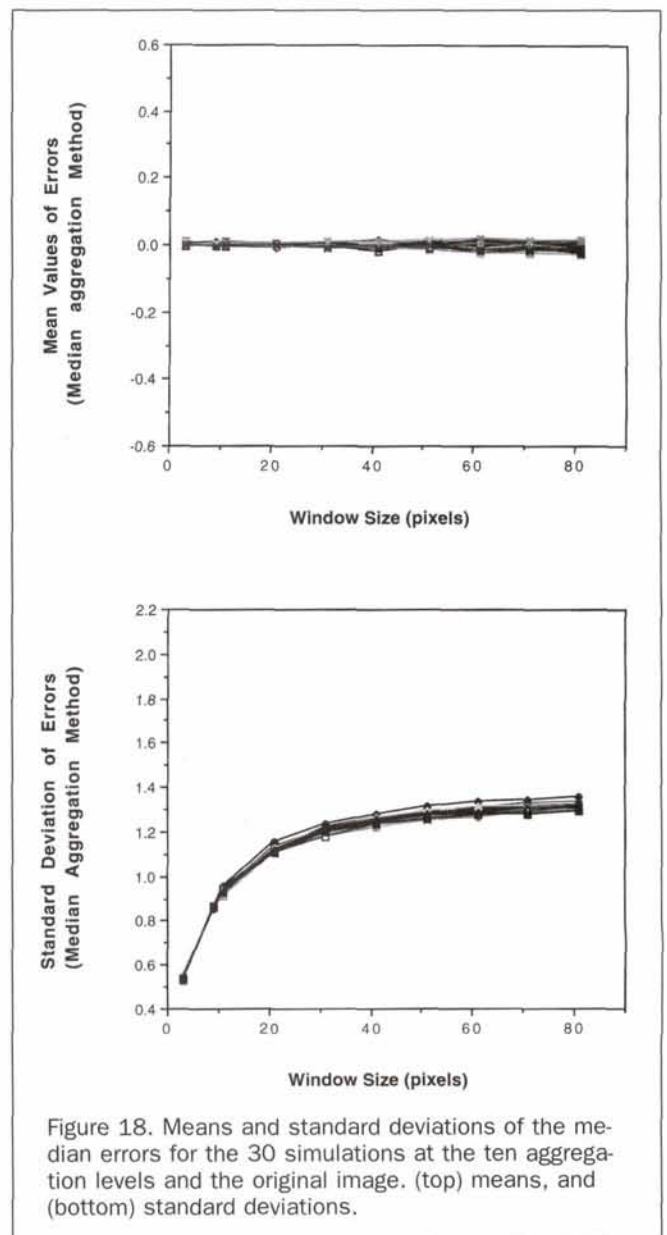

Figure 18. Means and standard deviations of the median errors for the 30 simulations at the ten aggregation levels and the original image. (top) means, and (bottom) standard deviations.
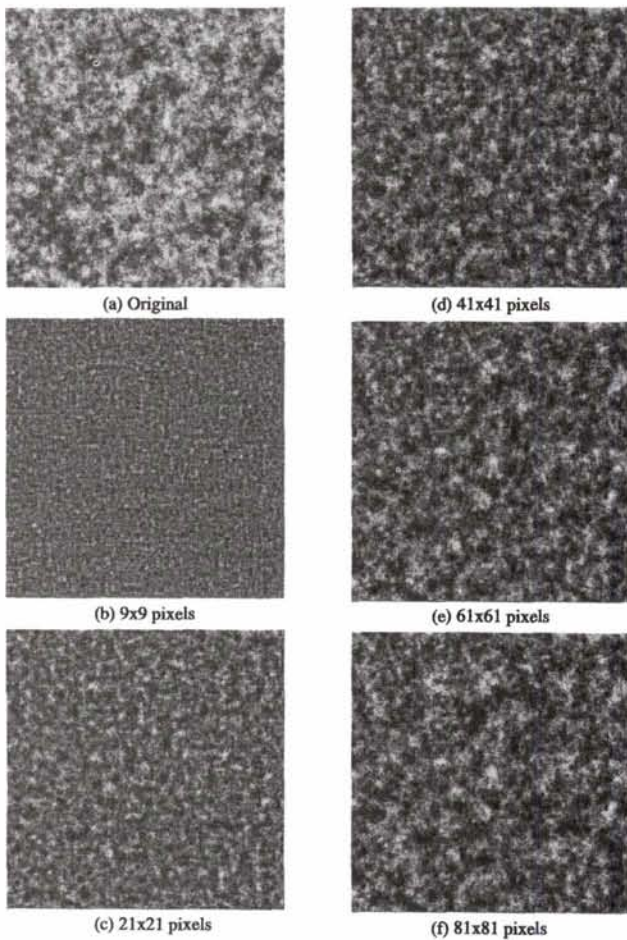
Figure 19. Images of the median errors at selected aggregation levels and the original image for one of the 30 simulations. (a) The original image, (b) 9- by 9-, (c) 21- by 21-, (d) 41- by 41-, (e) 61- by 61-, and (f) 81- by 81-pixel window aggregation.

Gupta, V.K., I. Rodriguez-Iturbe, and E.F. Wood, 1986. *Scale Problems in Hydrology*, D. Reidel Publishing Company, Boston, 245 p.

Hunter, G.J., and M.F. Goodchild, 1997. Modeling the uncertainty of slope and aspect estimates derived from spatial databases, *Geographical Analysis*, 29(1):35–49.

Heuvelink, G.B.M., 1992. An interactive method for multidimensional simulation with nearest neighbor models, *Proceedings of the 2nd CODATA Conference on Geomathematics and Geostatistics*, Nancy, France, pp. 51–57.

Lam, N., and D.A. Quattrochi, 1992. On the issues of scale, resolution, and fractal analysis in the mapping sciences, *Professional Geographer*, 44(1):88–98.

Lammers, R.B., L.E. Band, and C.L. Tague, 1997. Scaling behavior of watershed processes, *Scaling-Up* (P. van Gardinger, G. Foody, and P. Curran, editors), Cambridge University Press, Cambridge, UK, pp. 295–317.

Levin, S.A., 1993. Concepts of scale at the local level, *Scaling Physiological Processes, Leaf to Globe* (J.R. Ebleringer and C.B. Field, editors), Academic Press, Inc., New York, pp. 7–19.

Mark, D.M., and P.B. Aronson, 1984. Scale dependent fractal dimensions of topographic surfaces: An empirical investigation with applications in geomorphology and computer mapping, *Mathematical Geology*, 16(7):671–683.

Nellis, M.D., and J.M. Briggs, 1989. The effect of spatial scale on Konza landscape classification using textural analysis, *Landscape Ecology*, 2(2):93–100.

Peuquet, D.J., 1988. Representation of geographic space: toward a conceptual synthesis, *Annals of the Association of American Geographers*, 78(3):375–394.

Quattrochi, D.A., and M.F. Goodchild, 1997. Scale, multiscaling, remote sensing, and GIS, *Scaling of Remote Sensing Data for GIS* (D.A. Quattrochi and M.F. Goodchild, editors), Lewis Publishers, New York, pp. 1–11.

Seyfried, M.S., and B.P. Wilcox, 1995. Scale and the nature of spatial variability: Field examples having implications for hydrologic modeling, *Water Resources Research*, 31(1):173–184.

Stoms, D., 1992. Effects of habitat map generalization in biodiversity assessment, *Photogrammetric Engineering & Remote Sensing*, 58(11):1587–1591.

Turner, M.G., R.V. O'Neill, R.H. Gardner, and B.T. Milne, 1989. Effects of changing spatial scale on the analysis of landscape pattern, *Landscape Ecology*, 3(3/4):153–162.

Vieux, B.E., 1993. DEM Aggregation and smoothing effects on surface runoff modeling, *Journal of Computing in Civil Engineering*, 7(3):310–338.

Walsh, S.J., 1989. User consideration in landscape characterization, *The Accuracy of Spatial Databases* (M.F. Goodchild and S. Gopal, editors), Taylor & Francis, London, pp. 35–43.

Wolock, D.M., and C.V. Price, 1994. Effects of digital elevation model map scale and data resolution on a topography-based watershed model, *Water Resources Research*, 30(11):3041–3052.

Zhang, W., and D.R. Montgomery, 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations, *Water Resources Research*, 30(4):1019–1028.
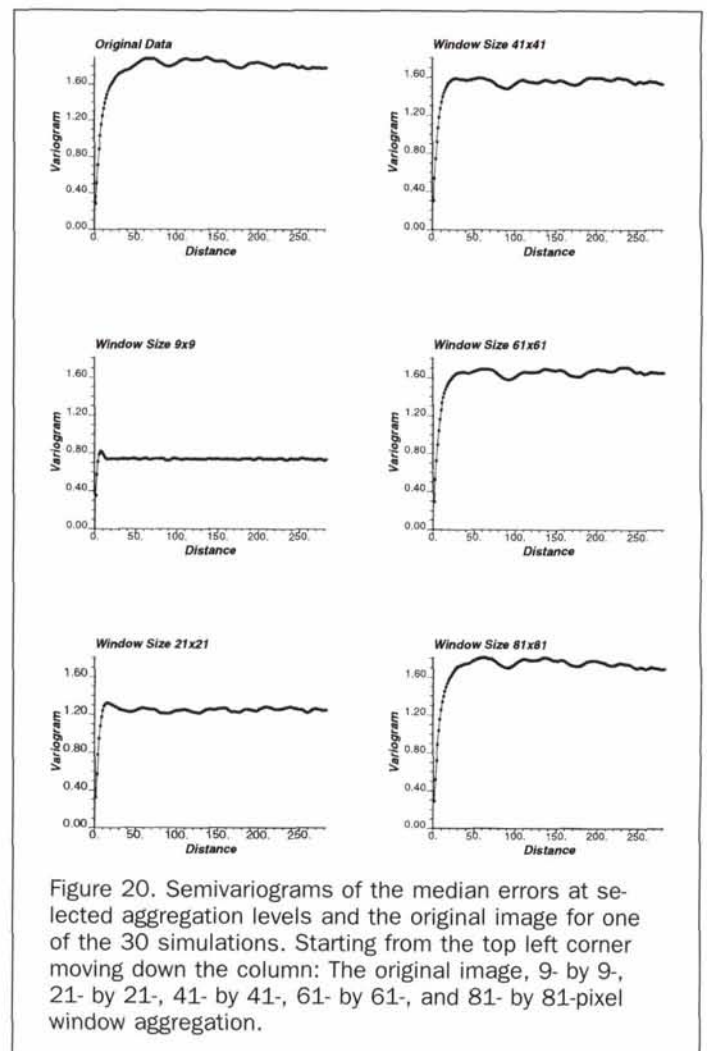
Figure 20. Semivariograms of the median errors at selected aggregation levels and the original image for one of the 30 simulations. Starting from the top left corner moving down the column: The original image, 9- by 9-, 21- by 21-, 41- by 41-, 61- by 61-, and 81- by 81-pixel window aggregation.