Gary J. Hunter
and Michael F. Goodchild

Modeling the Uncertainty of Slope and Aspect Estimates Derived from Spatial Databases

Estimates of slope and aspect are commonly made from digital elevation models (DEMs), and are subject to the uncertainty present in such models. We show that errors in slope and aspect depend on the spatial structure of DEM errors. We propose a general-purpose model of DEM errors in which a spatially autoregressive random field is added as a disturbance term to elevations. In addition, we propose a general procedure for propagating such errors through GIS operations. In the absence of explicit information on the spatial structure of DEM errors, we demonstrate the potential utility of a worst-case analysis. A series of simulations are used to make general observations about the nature and severity of slope and aspect errors.

1. INTRODUCTION

In recent years, there has been a dramatic increase in the use of digital computers to capture, store, process, analyze, and model geographic information, that is, information about specified locations on the Earth's surface. Much of this activity is subsumed under the rubric of geographic information systems (GIS; Burrough 1986; Maguire, Goodchild, and Rhind 1991), though it also owes much to related geographic information technologies of remote sensing and global positioning. It is now routine to create digital databases of such geographic variables as ground elevation, soil class, ownership, or land cover class, and to use them in analyses, and ultimately in decisions.

Although digital computers are relatively precise, the spatial databases used in GIS analyses are often of surprisingly low quality. Many are captured from printed maps, which are often designed to convey impressions about the distribution of geographic phenomena rather than precise scientific measurements. Many sets of data are the result of subjective interpretation of the landscape, and not replicable between observers. Thus the subject of uncertainty in spatial

The National Center for Geographic Information and Analysis is supported by the National Science Foundation under cooperative agreement SBR 88-10917.

Gary J. Hunter is professor of geomatics at the University of Melbourne. Michael F. Goodchild is professor of geography in the National Center for Geographic Information and Analysis, University of California, Santa Barbara.

Geographical Analysis, Vol. 29, No. 1 (January 1997) © 1997 Ohio State University Press
Submitted: 1/17/96. Revised version accepted: 7/11/96.

databases has received increasing prominence as users of GIS have come to realize that (1) the quality (or fitness for use) of their products may be less than that required for their tasks; and (2) that without knowledge of quality, the integrity of past, present, and future decisions may be jeopardized. The latter concern is particularly pertinent to agencies responsible for making regulatory decisions that may be subject to judicial review.

In this paper we use the term "uncertainty" to denote a lack of knowledge of true value, that is, the value that would be discovered if one were to visit the field and make an observation using a perfectly accurate instrument. Used in this way, uncertainty includes "error," or uncertainties due to imperfections in measuring instruments, as well as the effects of cartographic generalization, which alter observations in the interests of cartographic simplicity and related objectives. It also includes uncertainty due to inadequate definition, and consequent variation between observers. To be consistent with usage in statistics, we use the term "error" in this paper, in the context of "error model," in the broader sense of uncertainty rather than the narrow sense of mistake. We also use the term "error" to refer to an instance of the difference between observed or recorded value and the corresponding true value.

For the purposes of this paper we adopt a very simple model of the operation of a GIS. Observations and measurements are used to populate the GIS database, and are subject to uncertainty. The functions of the GIS are used to process and analyze the data, perhaps incorporating them as input to models. Finally, the results of processing are output as products, and form the basis for decisions. In this simple model, three steps are necessary to address the issue of quality: (1) develop error models for each data component; (2) determine how uncertainty propagates during product creation; and (3) devise appropriate measures and visualizations of product uncertainty. This process is particularly important when the response of output to input is nonlinear, or when it is otherwise difficult to anticipate intuitively the effects of input uncertainty on output.

The theory of error analysis is well developed for scalar measurements (Taylor 1982), and relies heavily on the Gaussian distribution. In surveying, the theory of errors in point positions is based on the multivariate Gaussian distribution and forms the basis of least-squares adjustment (for example, Leick 1990, pp. 105–26). But in many spatial databases it is impossible to identify the original measurements that support each item of information, since much of a data set may have been obtained by interpolation between observations. Such complex patterns of lineage lead to strong interdependencies between the errors in items of information, particularly when the items are located close together in space.

The need to develop effective error models for spatial databases is reflected in a growing literature (Fisher 1991; Lee, Snyder, and Fisher 1992; Goodchild and Gopal 1989; Goodchild, Sun, and Yang 1992; Heuvelink 1993). These models form the first stage of the approach described above. In some cases it is possible to approach the propagation of uncertainty analytically as in classical error analysis [for example, Heuvelink, Burrough, and Stein (1989) use a Taylor series expansion] but in other cases it is necessary to adopt a Monte Carlo approach (Openshaw, Carver, and Carver 1991). If a suitable stochastic process can be devised to model errors, then a sample of realizations of the process is generated; the analysis is performed repeatedly on each realization; and finally product uncertainty is computed by evaluating some suitable statistical summary of the range of outputs, such as a standard error. Because of the strong interdependencies of errors, each realization in this approach must be a complete data set, or "map." Thus we can define an error model for a spatial data set as a stochastic process capable of generating a population of data sets, such that the differences between data sets are representative of the uncertainty present in each. In essence, we have substituted "data set" or "map" for the individual scalar measurement of the classical Gaussian error model; one might therefore express the objective of this research as "finding a Gaussian distribution for maps."

The elevation of the land surface is a particularly well-defined geographic variable (problems of definition do nevertheless exist, but will not be reviewed here), and data sets capturing elevations over defined geographic areas are widely used for purposes ranging from hydrologic modeling to the siting of transmitters for cellular telephone networks. Of the various alternative data models, the most widely available is the digital elevation model, or DEM, which consists of a rectangular grid of sample elevations. DEMs with a grid spacing of three seconds of arc are widely available, and extensive areas of the U.S. land surface are covered by available DEMs with a thirty-meter spacing. However, like much geographic data, DEMs present problems for error modeling because of the complex pattern of dependencies between observed elevations at grid sample points, and the data used to create the DEM. Several different methods are regularly used, including manual photogrammetric transects, automated photogrammetry, and interpolation between contours.

Slope and aspect are important terrain parameters for many types of environmental applications. Several authors have investigated the error effects of the algorithms used to calculate these quantities from DEMs, including Carter (1990), Skidmore (1989), Smith, Prisley, and Weih (1991), Srinivasan and Engel (1991), and Wood and Fisher (1993). Carter (1992) has looked at the effect of rounding elevations to the nearest meter (standard practice with the USGS data used here), while Isaacson and Ripple (1990) have examined the implications that different cell sizes have for the two quantities, and Kumler (1990) has investigated the error effects of the various methods available for creating DEMs. While this research has shown the relative merits (or otherwise) of many of the algorithms available, apart from the work of Carter (1990) it has not dealt with the effects of elevation error as an input to those algorithms.

Hunter and Goodchild (1995) presented an error model for DEMs. The observed value is assumed to be the sum of a true value and a disturbance term; the latter will be modeled in this paper as a spatially correlated random field. They showed how the model could be used to visualize the uncertainty in the position of a contour on a topographic map. In this paper we explore the implications of the model for analyses of DEMs, particularly the calculation of slope and aspect. We examine the propagation of data error through these two common GIS operations, and the influence of the pattern of dependencies between errors on the results. We discuss practical strategies for coping with lack of knowledge of the model's spatial dependence parameter ρ. The approach is partly analytic, and partly numeric. The paper is structured so as to provide an overview of DEM error and testing procedures in section 2, followed by a description of the test data set in section 3, and discussion of the model to be applied in section 4. Spatial dependence is introduced in section 5, together with a detailed description of the steps involved in implementing the model. Finally, section 6 presents the model's application to assessing uncertainty in slope and aspect, while section 7 discusses the results and implications.

2. GENERAL DISCUSSION OF DEMS

DEMs are digital representations of terrain surfaces in which elevations are recorded over a rectangular grid of points. Often the grid values will have been interpolated or resampled from other, possibly irregularly spaced sample points (Burrough 1986; Skidmore 1989), and in addition the process of creating the DEM will have introduced a number of additional errors, some more systematic than others. For instance, when aerial photography is photogrammetrically profiled along transects, operators controlling the process tend to underestimate elevation when moving uphill and overestimate it when moving downhill. This produces a characteristic "Firth Effect" whereby adjacent cells in the same transect exhibit high positive correlation of errors while neighboring cells in adjacent transects show high negative correlation of errors. The pattern becomes obvious when the DEM is displayed by interpolating contours.

In cases where the GPM2 automated image correlation system has been employed, the DEM is constructed in 9 × 8-millimeter patches (at photography scale), and the assembled DEM may show distinct steps in elevation (up to ten meters in magnitude from the authors' experience) at patch boundaries throughout the model [see Hunter and Goodchild (1995) for an illustration of this effect]. In addition, there is a rounding error associated with providing all elevation values as integers. These are just some of the errors which may occur in DEMs and, while it is not the purpose of this paper to present a comprehensive review, with the numerous combinations of data sources and processing methods employed to produce DEMs there is clearly much that remains unknown about this topic. Carter (1988, 1989) and Theobald (1989) are suggested for further reading.

The U.S. Geological Survey (USGS) is a major producer of DEMs and since this research uses one of their products it is worth examining their error assessment techniques. Considering only their 7.5-minute products, their method (described in USGS 1990) is to test the DEM by comparing a minimum of twenty-eight test points (eight on the boundary and twenty inside the file) with ground control values that are taken to represent true elevations. The root mean square error (RMSE) is the resultant summary statistic provided with the file, defined as follows:

$$RMSE = \left[\sum_{i=1}^{n} (z_i - z^*_i)^2 / n \right]^{1/2} \tag{1}$$

where n is the number of sample points, i denotes a sampled point, z_i is the elevation of the DEM at the sampled point, and z^*_i is the true value at the sampled point.

Level 1 DEMs (the most common) are generally considered to have an RMSE of less than or equal to seven meters, although some may have an RMSE as high as fifteen meters, and the stated USGS policy in this regard is that

> an absolute elevation error tolerance of fifty meters (approximately three times the fifteen-meter RMSE) be set for blunders for any grid node when compared to the true height from mean sea level ... (USGS 1990, p. 14)

While not specifically mentioned in their documentation, the USGS has clearly assumed a Gaussian distribution of the elevation error and the authors have made the same assumption in this paper. Level 2 and Level 3 DEMs are considered more accurate and have more stringent tolerances.

The focus of this paper is on the derivation of slope and aspect estimates from DEMs. To avoid possible ambiguity, we define gradient as a vector quantity with components equal to the partial derivatives of the surface in the x and

y directions. Slope S is defined as the magnitude of this vector, or the tangent of the angle of steepest slope of a plane tangential to the surface (we use the term slope angle to refer to this quantity's arc tangent). Aspect A is defined as the direction of the horizontal projection of the line of steepest slope, or the arc tangent of the negative ratio of the gradient vector's two components. Thus:

$$S = [(\partial Z/\partial x)^2 + (\partial Z/\partial y)^2]^{1/2} \tag{2}$$

and

$$\tan A = \frac{-\partial Z/\partial y}{\partial Z/\partial x}. \tag{3}$$

3. THE TEST SITE

The data for the exercise comes from a 448 row by 334 column subset (86 percent) of the USGS DEM for the 7.5-minute, 1:24,000 State College (Pennsylvania) mapping quadrangle. It should be noted that the scale of 1:24,000 has no relationship to the DEM, other than to assist in identifying the map quadrangle for which the DEM applies. The total number of elevation values in the test file is 149,632. The values represent elevations of points in a rectangular array defined on a projection of the Earth's surface with a spacing of thirty meters.

The DEM was supplied as part of the "Visualization of Spatial Data Quality Challenge"—a research contest jointly sponsored by the U.S. National Center for Geographic Information and Analysis (NCGIA); the U.S. Environmental Protection Agency (EPA) Center for Environmental Statistics; the U.S. Department of Agriculture Soil Conservation Service (SCS); and the Statistical Graphics Section of the American Statistical Association (Beard 1992). Figure 1 shows a contour plot of the test site topography at a contour interval of twenty meters. The test site measures approximately ten kilometers by thirteen kilometers and has considerable variation in terrain, with elevation values ranging from 255 meters in the north to 743 meters in the southeast.

We assume for the purposes of this paper that as a Level 1 DEM the test data set has a spatially stationary RMSE of seven meters. In doing so we may be overestimating the actual uncertainty, since seven meters is a maximum for this class of DEMs, not a measured value for this instance; and we may also be ignoring correlations between RMSE and other factors—it may be, for example, that areas of steeper slope have larger errors.

4. THE ERROR MODEL

The Gaussian model (where the mean of the population is an estimate of the true value and the standard deviation is a measure of the uncertainty of the observations) is commonly used as an error model, but it makes only the most general assumptions about the processes by which the error has accumulated (total error is assumed to be the sum of a large number of random, additive effects). When additional information is available about the structure of errors in a data set, the Gaussian model should be replaced with a more accurate representation.

We use a similar approach here, by proposing that in the absence of more definitive information, error in a DEM can be modeled by adding to the true

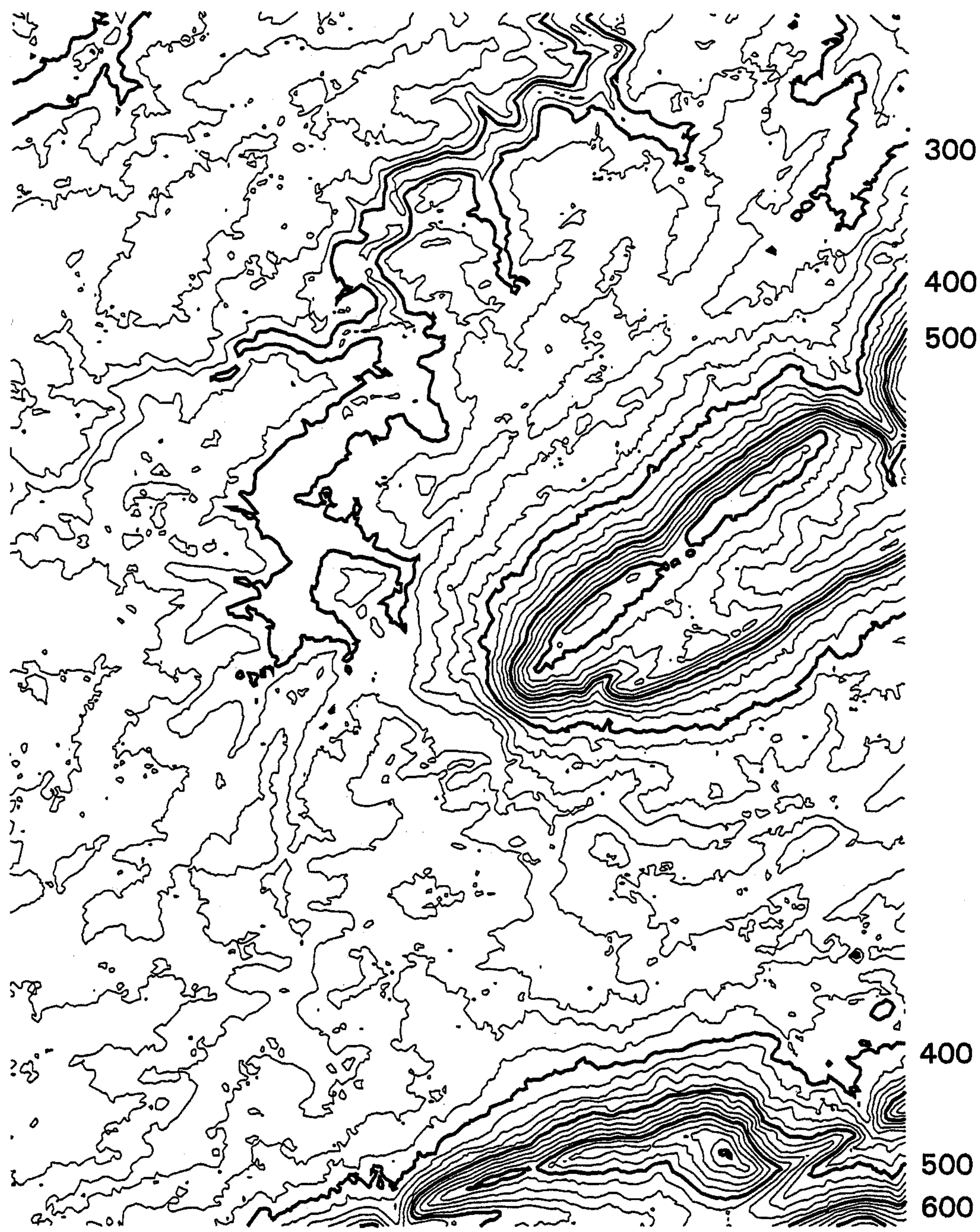

FIG. 1. Test Site Topography (shown at a contour interval of 20m, with 100m contours labeled)

value a disturbance term consisting of a spatially autocorrelated random field, with parameters that are stationary across the project area. When these assumptions are known to be invalid, the model should be modified accordingly, by substituting a more complex pattern of spatial dependence of errors, or a multiplicative error term, or spatially varying parameters, as appropriate.

A variety of methods can be used to estimate slope and aspect from a DEM grid. Srinavasan and Engel (1991) review four such methods: weighted least squares fit of a plane to a 3×3 neighborhood centered on each point (Burrough 1986, p. 50); exact fit of a nine-term partial quartic surface to a 3×3

neighborhood (Zevenbergen and Thorne 1987); finding the maximum slope between a point and its eight queen's-case neighbors (Shanholtz et al. 1990); and a method due to Beasley and Huggins (1982). Of these four methods, the first is most amenable to a mathematical analysis of error propagation.

Consider a 3×3 neighborhood centered on (0,0) and with spacing 1 between points. If rook's-case neighbors are given weights of 2 and bishop's-case neighbors weights of 1, reflecting the former's greater proximity, the least-squares estimators for the x and y components of the gradient vector are

$$\partial Z/\partial x = [z_{1,1} + 2Z_{1,0} + Z_{1,-1} - Z_{-1,1} - 2Z_{-1,0} - Z_{-1,-1}]/8 \tag{4}$$

$$\partial Z/\partial y = [Z_{1,1} + 2Z_{0,1} + Z_{-1,1} - Z_{1,-1} - 2Z_{0,-1} - Z_{-1,-1}]/8. \tag{5}$$

Now assume that the RMSE of each elevation error is s_z, and that the correlation between errors at pairs of grid points is a function only of the distance between them. Because the estimating equations are linear, the RMSEs of the components of the gradient vector, s_x and s_y, can be easily obtained as

$$s_x^2 = s_y^2 = s_z^2[6 + 8r(1) - 4r(2) - 8r(\sqrt{5}) - 2r(\sqrt{8})]/32 \tag{6}$$

where $r(d)$ denotes the (signed) correlation between elevation errors at points distance d apart.

The correlation between errors in the two components of the gradient vector is readily shown to be 0, despite the existence of common terms in the estimating expressions. If errors in grid point elevations are also independent ($r = 0$), then $s_x = s_y = 0.433 s_z$ (for example, a seven meters RMSE and a grid spacing of thirty meters will produce an RMSE of 0.101 in estimates of the components of the gradient vector; on a slope angle of 20 degrees the addition of two RMSE is equivalent to a slope angle of 29.5 degrees). If errors are perfectly correlated ($r = 1$), then as expected $s_x = s_y = 0$, and there is no uncertainty in estimates of slope or aspect. Unfortunately expressions for error in slope and aspect are not as simple as those for the components of the gradient vector because of the nonlinear nature of the estimating equations.

Thus a more useful approach to error modeling might be to bypass the analytic treatment of individual algorithms entirely, by examining the effects that changes in input parameters have upon the outputs via a simulation process. By using such a model, an alternative means can be provided for overcoming the first two requirements of the three-part error management approach given in section 1, which in turn provides the basis for considering the third component of managing uncertainty—visualization (or communication) of the results to users.

Although the correlation function $r(d)$ clearly has a critical effect on the uncertainty of slope and aspect calculations, and although estimates of RMSE for published DEMs are generally available, almost nothing is known about the empirical spatial structure of DEM errors, or how it varies between methods for DEM generation. We explore this critical issue of spatial dependence in the next section.

5. DISCUSSION OF SPATIAL DEPENDENCE

5.1 *Case 1: Spatial Independence* ($r = 0$)

The first option available is to consider the error in the elevation of each point to be spatially independent of its neighbors ($r = 0$). In other words, knowledge of the error present at one point provides no information on the errors present at neighboring points, even though the elevations themselves may have similar values. Following the definition of an error model given in section 4, a realization can be achieved by disturbing each elevation by an independent disturbance term:

$$z(x,y) = z^*(x,y) + N(0,s_z) \tag{7}$$

where $z^*(x,y)$ is the observed elevation at (x,y) and $N(0,s_z)$ is an independent, normally distributed random variable with mean 0 and standard deviation s_z. Heuvelink (1993) illustrated the impact of this spatially independent model on slope and aspect estimates using Monte Carlo simulation, and by a Taylor series expansion of the relevant equations.

In this analysis we have assumed that the mean error is zero, and thus that the RMSE statistic is equal to the standard deviation (s_z) of error. Monckton (1994) and Li (1991) discuss the presumption by DEM producers that there is no systematic bias in DEM elevations, and note evidence that small nonzero mean errors for DEMs exist, particularly over certain regions of a map.

Figure 2 shows a realization of the model as a contour map. Addition of a spatially independent disturbance term has raised and lowered adjacent grid elevations and created a landscape that is clearly much more rugged and noisy than the observed landscape shown in Figure 1. The uncertainties of the two components of the gradient vector are also very high, implying that slope and aspect estimates from such data are highly unreliable. On visual appearance alone, it is clear that this model of independent disturbances is not acceptable. Thus although no information about spatial dependence of errors is available from the producers of this DEM, we can safely surmise that correlations between adjacent errors are in reality positive. Indeed, if they were not, the practice of estimating slope and aspect from published DEMs would be close to useless.

5.2 *Case 2: Spatial Dependence (limit $r = 1$)*

At the other extreme, consider the limit where spatial autocorrelation is maximum, all errors are perfectly correlated, and there is in effect only 1 degree of freedom in the disturbance field being applied to the DEM. This is equivalent to specifying a systematic error in elevations, or a mean error not equal to zero; it seems unlikely that any DEM production process would generate errors of this nature. On the other hand, it is possible to imagine appropriate motivating circumstances. Suppose an area is to be flooded, but the eventual elevation of the water surface is uncertain. We know, however, that the elevation is perfectly correlated in space, and that the water surface therefore has only one degree of freedom.

5.3 *Case 3: Spatial Dependence* $(0 < r < 1)$

The case of positive correlation less than 1 is clearly most realistic. Of the many methods available for defining the disturbance field $e(x,y)$, we adopt the spatially autoregressive process:

$$\mathbf{e} = \rho\mathbf{W}\mathbf{e} + \mathbf{N}(0,1) \tag{8}$$

where $\mathbf{e}$ denotes a vector of grid values of the disturbance field, ρ is a parameter, $\mathbf{W}$ is a matrix of weights, and $\mathbf{N}(0,1)$ is a vector of independently and normally distributed random deviates. The elements of the weights matrix are set to 1 for

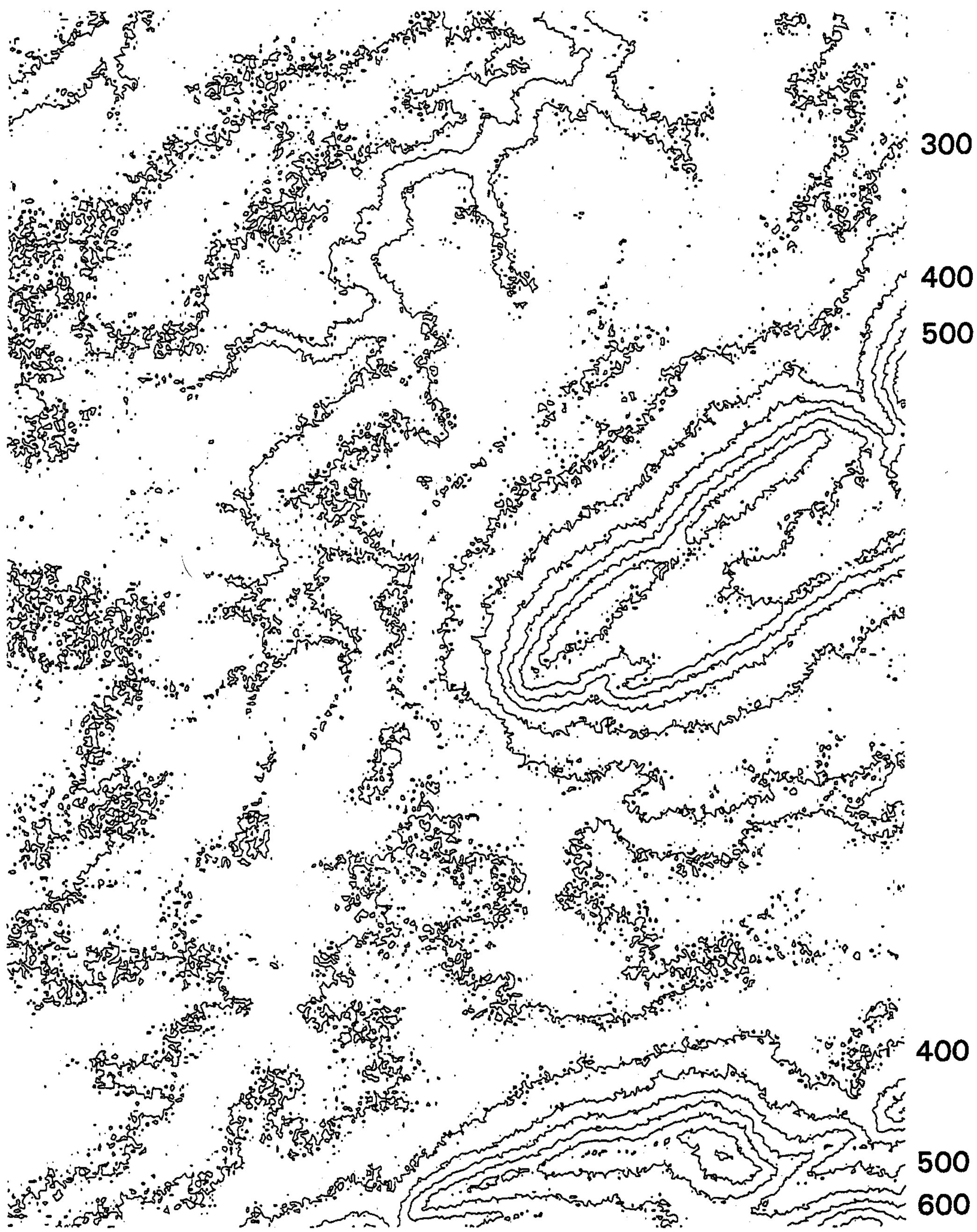

FIG. 2. Topography Modified by a Spatially Independent Disturbance Term [Equation (7); shown at a contour interval of 50m]

rook's-case neighbors and zero otherwise. This definition of the weights matrix forces ρ to lie in the range 0 to 0.25.

To simulate disturbance fields, we use the iterative approach proposed by Heuvelink (1992), which computes left-hand-side values of **e** and inserts them on the right-hand side until convergence occurs.

To illustrate, Figure 3 shows six realizations ($\rho = 0.0000$, 0.2000, 0.2400, 0.2450, 0.2490, and 0.2499) to demonstrate the effect of increasing ρ. Note that in the range $0 < \rho < 0.20$ there is little obvious change in the autocorrela-

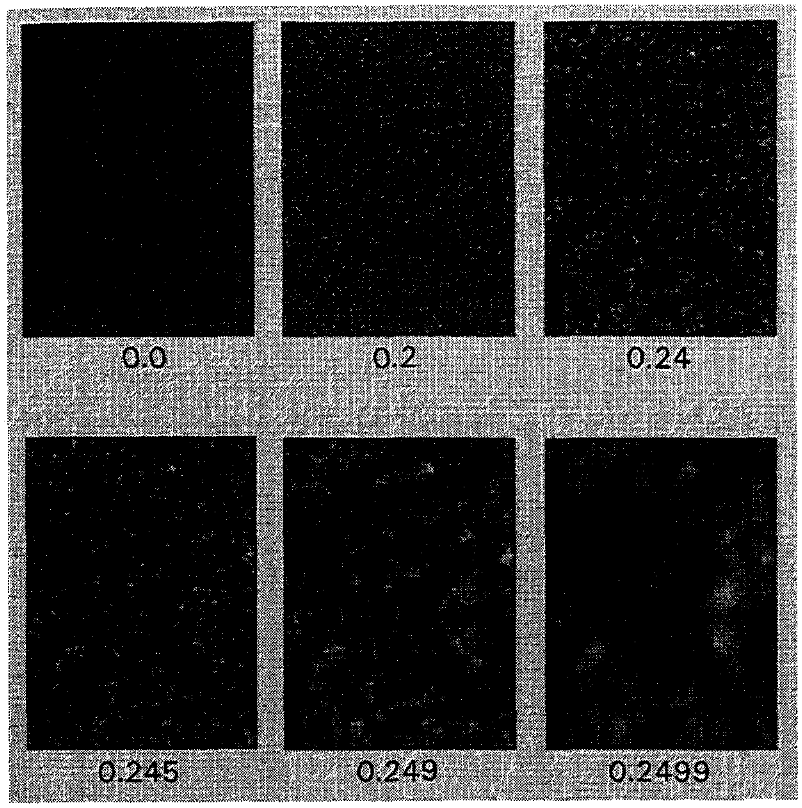

FIG. 3. Realizations of the Spatially Autoregressive Disturbance Field for $\rho = 0.0000$, 0.2000, 0.2400, 0.2450, 0.2490, and 0.2499, Using the Same Random Number Seed. Cells are shaded using a 256-value gray scale.

tion effect, and it is not until ρ takes values very close to the limit of 0.25 that distinct patterns start to emerge.

In principle, the correlation function $r(d)$ of realizations of this autoregressive process depends only on ρ, and equation (6) could thus be rewritten to express the standard errors s_x and s_y as functions of ρ also. Unfortunately while it is clear that $r(d)$ is a decreasing function, it has not been possible to obtain a simple closed-form expression in this case (Whittle 1950). Thus for the remainder of the paper we discuss results in terms of the parameter ρ, and suggest that correlograms and standard errors be obtained by simulation.

6. APPLICATION OF THE MODEL TO SLOPE AND ASPECT ESTIMATES

In order to investigate the effect of simulated changes in elevation at different levels of spatial autocorrelation on calculations of slope and aspect, an experiment was designed whereby ten realizations of the error model for the test site DEM were made for each value of ρ from 0.0000 to 0.2000 (in 0.0100 increments), from 0.2000 to 0.2400 (in 0.0050 increments), from 0.2400 to 0.2490 (in 0.0010 increments), and for 0.2495 and 0.2499. These simulated error fields were each added to the DEM, and slope and aspect estimates derived using the 3×3 neighborhood least squares method analyzed previously. The process is summarized in Figure 4.

Calculating the difference between "true" and realized aspect values requires that the angular differences be interpreted correctly. For example, the difference between aspects of 5 degrees and 355 degrees is 10, not 350.

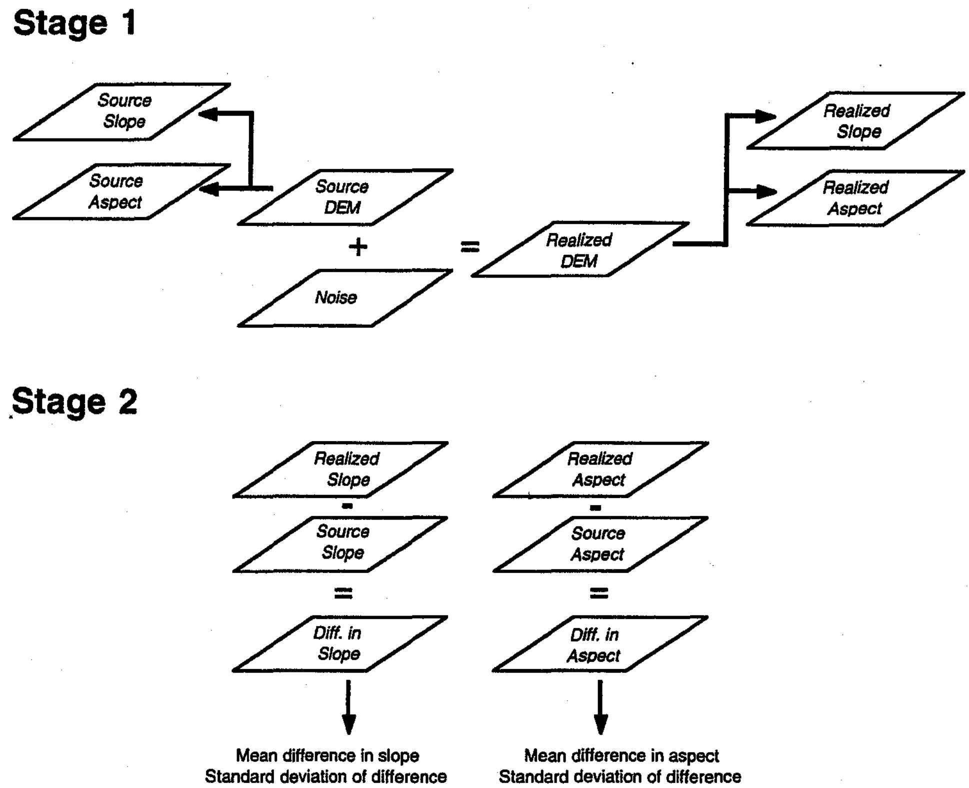

FIG. 4. The Two-Stage Procedure for Deriving Realized Slope and Aspect Grids, Followed by Determination of Differences between "True" Grid and "Distorted" Grid for the Two Quantities

The first assessment of the experimental results is shown in Figure 5, which shows the effect of the realization procedure on slopes. The cells in each image have been shaded according to their slope angle (with white representing 0 degrees slope angle and black representing 45 degrees, the maximum value occurring in any of the data sets). The slope estimates for the source DEM are given in the lower-right image. Note the increase in feature clarity with increasing ρ, as expected. Distinctive features such as the river valley in the top center of the test site and the mountain ranges in the east and southeast remain relatively unaffected due to their size in relation to the RMSE of 7m. The parallel linear features visible near the lower edge of some images, particularly the source DEM, are examples of the GPM2 processing error discussed in section 2.

Figures 6 and 7 show statistical summaries of the differences in slope and aspect between the source DEM and error model realizations. Figure 6 shows standard deviations of differences in slope angle plotted against ρ; Figure 7 shows the same results for aspect. Although results are shown for all four hundred simulations, there is little variation among the realizations for each value of ρ in either figure.

In Figure 6, standard deviations are generally steady up to $\rho = 0.2000$, then there is a transition zone between $\rho = 0.2000$ and 0.2400, followed by a sharp decrease to almost zero. In Figure 7 there is a more distinct turning point at around $\rho = 0.2400$. The implications of these results are discussed further in section 7.

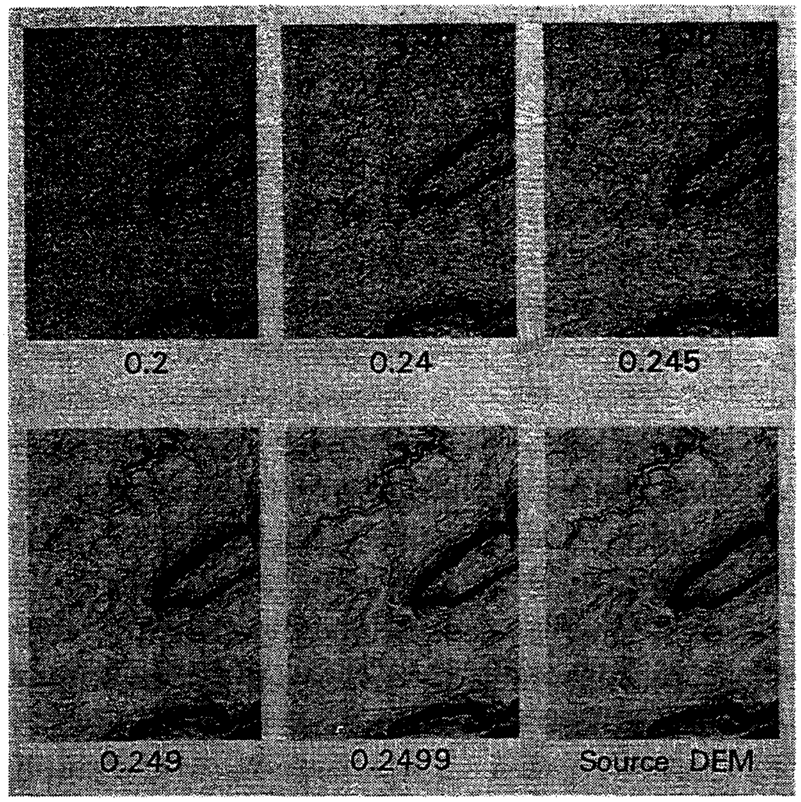

FIG. 5. Realizations of Slope for the DEM with $\rho = 0.2000$, 0:2400, 0.2450, 0.2490, and 0.2499, Using the Same Random Number Seed. Cells have been shaded by slope angle, with white = 0 degrees and black = 45 degrees.

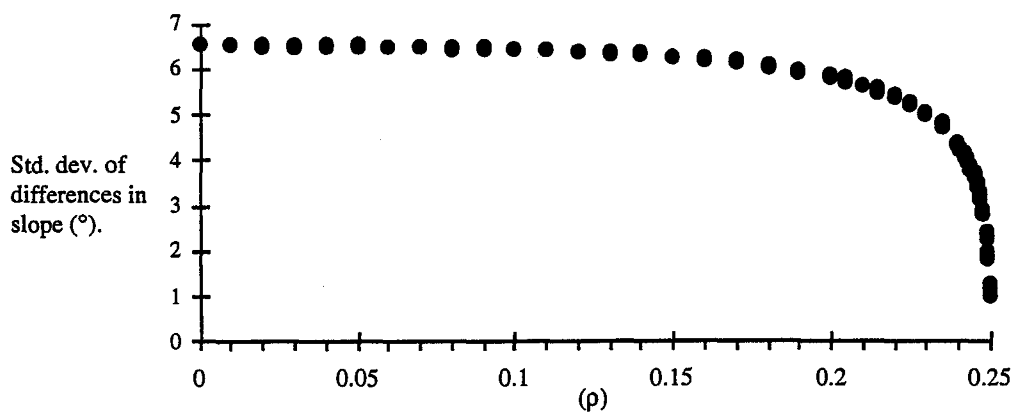

FIG. 6. Standard Deviation of Differences in Slope Estimates Plotted against ρ. Ten replications were made for each value of ρ.

7. DISCUSSION OF RESULTS

The results obtained from the experiment and their operational application are now discussed. Firstly, the model provides a means of creating different realizations of an output that is software and hardware independent, since it operates on whatever algorithms and processor are being used at the time. Even where an empirical error model can be defined or the effects of error propagation obtained by analysis, the user is not necessarily able to understand

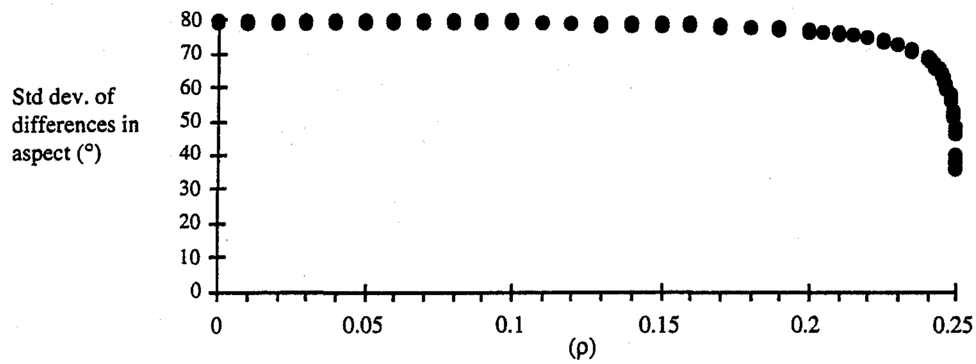

FIG. 7. Standard Deviation of Differences in Aspect Estimates Plotted against ρ.

what the final variation in product may be. The general approach given in this paper overcomes that problem.

Lack of knowledge of the structure of spatial dependence in DEM errors presents a serious problem, given the critical importance of this structure in determining uncertainty in the products of many GIS calculations. Since it seems unlikely that such information will become routinely available for published DEMs, at least in the near future, we propose an alternative strategy that may still provide useful results. Study of the graphs in the preceding section shows that there is a point in the domain $0 < \rho < 0.25$ below which there is no significant change in the difference between source and realized DEMs. This could be selected as a worst-case ρ value, and a series of realizations might be made at this level representing the outer limit of variation in aspect that a user could expect from the specific combination of DEM, its RMSE, and the aspect algorithm employed. The transition point for slope is not as well defined as it is for aspect, but a user could still apply a minimum ρ value of about 0.2 (Figure 6) with confidence that the mean difference in slope angle from the process will not increase by more than about 1 degree (which is within the bounds of the data set used, given that the combination of thirty meters grid cells with elevations rounded to the nearest meter produces a precision in slope calculation of no better than about 1.9). Indeed, it might not be necessary to look at the graphs to assess where the transition value of ρ lies, since users may be able to judge its value effectively by examining the output. This can be seen in Figure 5, where realizations at $\rho = 0.2000$ essentially remain featureless in flatter terrain due to their randomness, and it is only those features whose size is at least several times greater than the RMSE of the DEM that are maintained. Thus, presenting users with single, initial realizations at selected ρ values could just as easily allow them to choose the limit at which they wish to operate.

The procedure has widespread application to many operations. These include land suitability/capability analysis; soil classification where slopes are used to assist in class definition; viewshed calculations; computation of parameters such as "northness," which is used for assessing vegetation reflectance; derivation of combined elevation/slope/aspect parameters for vegetation growth indices; calculation of snow wetness when examining spring thaw runoffs; and throughout hydrologic modeling in general. In this paper, only slope and aspect realizations are shown, but there is no reason why each realization should not be further processed by whatever models and algorithms a user has applied to the source DEM, in order to give realizations of other derived quantities. Indeed, the process applies not just to DEMs but also to other attributes that may be modeled by grid cell structures, such as noise or pollution values. The procedure might also be used to test the effect of different DEM cell resolutions or RMSEs, in addition to alternative process models and algorithms. Finally, a series of realizations at selected ρ values might be produced and cell counts recorded to develop confidence levels in cells to satisfy nominated criteria (for example, whether they are "seen/not seen" in a viewshed, or "above/below 10 degree slope" for a land use study).

8. CONCLUSIONS

In this paper, a spatially autoregressive error model has been applied to study the effects of uncertainty in digital elevation models upon slope and aspect estimates derived from them. Use of the term "uncertainty" is preferred to "error," since it conveys the fact that it is our lack of empirical knowledge about errors in the source data and the algorithms used in the computational process that has led to a lack of understanding about the quality of the final product. The model can be described as a stochastic process capable of generating a population of distorted versions of the same reality, with each version being a sample from the same population. A procedure is given for developing alternative realizations of slope and aspect maps with differing levels of spatial autocorrelation (ρ). It has been demonstrated both analytically and empirically that errors in slope and aspect depend on the structure of spatial dependence of errors. Thus it is important that producers of DEMs develop appropriate methods for measuring and characterizing such structure, in addition to the traditional RMSE. Each method of DEM production likely has its own structure of spatial dependence, and the spatially autoregressive model used in this paper may not be ideal for any of them. Nevertheless, just as the Gaussian distribution is accepted as a first approximation to the distribution of errors in scalar measurements, the methods used in this paper remain appropriate as a first approximation until more is known about the real spatial structure of DEM errors.

LITERATURE CITED

Beard, M. K. (1992). "Call for Participation: Visualization of Spatial Data Quality Challenge." Santa Barbara, Calif.: National Center for Geographic Information and Analysis, University of California, 4 pp.

Beasley, D. B., and L. F. Huggins (1982). ANSWERS (areal nonpoint source watershed environmental response simulation): User's Manual. U.S. EPA-905/9-82-001, Chicago, Ill.: U.S. Environmental Protection Agency.

Burrough, P. A. (1986). *Principles of Geographic Information Systems for Land Resources Assessment*. Oxford: Clarendon Press.

Carter, J. R. (1988). "Digital Representations of Topographic Surfaces." *Photogrammetric Engineering and Remote Sensing* 54(11), 1577–80.

——— (1989). "Relative Errors Identified in USGS Gridded DEMs." *Proceedings of the Ninth International Symposium on Computer-Assisted Cartography (AUTO CARTO 9), Baltimore, Maryland*, pp. 255–65.

——— (1990). "Some Effects of Spatial Resolution in the Calculation of Slope Using the Spatial Derivative." *Proceedings of the 1990 ACSM-ASPRS Annual Convention, Denver, Colorado*, vol. 1, pp. 43–52.

——— (1992). "The Effect of Data Precision on the Calculation of Slope and Aspect Using Gridded DEMs." *Cartographica* 29(1), 22–34.

Fisher, P. F. (1991). "Modelling Soil Map-Unit Inclusions by Monte Carlo Simulation." *International Journal of Geographical Information Systems* 5(2), 193–208.

Goodchild, M. F., and S. Gopal, eds. (1989). *Accuracy of Spatial Databases*. London: Taylor and Francis.

Goodchild, M. F., G. Sun, and S. Yang (1992). "Development and Test of an Error Model for Categorical Data." *International Journal of Geographical Information Systems* 6(2), 87–104.

Heuvelink, G.B.M. (1992). "An Iterative Method for Multidimensional Simulation with Nearest Neighbor Models." *Proceedings of the 2nd CODATA Conference on Geomathematics and Geostatistics, Nancy, France*, pp. 51–57.

——— (1993). *Error Propagation in Quantitative Spatial Modelling: Applications in Geographical Information Systems*. Utrecht: Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht.

Heuvelink, G. B. M., P. A. Burrough, and A. Stein (1989). "Propagation of Errors in Spatial Modelling with GIS." *International Journal of Geographical Information Systems* 3(4), 303–22.

Hunter, G. J., and M. F. Goodchild (1995). "Dealing with Uncertainty in Spatial Databases: A Simple Case Study." *Photogrammetric Engineering and Remote Sensing* 61(5), 529–37.

Isaacson, D. L., and W. J. Ripple (1990). "Comparison of 7.5-minute and 1-degree Digital Elevations Models." *Photogrammetric Engineering and Remote Sensing* 56(11), 1523–27.

Kumler, M. P. (1990). "A Quantitative Comparison of Regular and Irregular Digital Terrain Models." *Proceedings of the GIS/LIS '90 Conference, Anaheim, California*, vol. 1, pp. 255–63.

Lee, J., P. K. Snyder, and P. F. Fisher (1992). "Modeling the Effect of Data Errors on Feature Extraction from Digital Elevation Models." *Photogrammetric Engineering and Remote Sensing* 58(10), 1461–67.

Leick, A. 1990. *GPS Satellite Surveying*. New York: Wiley.

Li, Z. 1991. "Effects of Check Points on the Reliability of DTM Accuracy Estimates Obtained from Experimental Tests." *Photogrammetric Engineering and Remote Sensing* 57(10), 1333–40.

Maguire, D. J., M. F. Goodchild, and D. W. Rhind (1991). *Geographical Information Systems: Principles and Applications*. London: Longman Scientific and Technical.

Monckton, C. G. (1994). "An Investigation into the Spatial Structure of Error in Digital Elevation Data." In *Innovations in GIS 1*, edited by M. Worboys, pp. 201–11. London: Taylor and Francis.

NCGIA (1990). "Unit 46-Managing Error." In *Core Curriculum: Technical Issues in GIS*, vol. 2, edited by M. F. Goodchild and K. K. Kemp, pp. 46.3–46.10. Santa Barbara, Calif.: National Center for Geographic Information and Analysis.

Openshaw, S., M. Carver, and S. Carver (1991). "Error Propagation: A Monte Carlo Simulation." In *Handling Geographic Information: Methodology and Potential Applications*, edited by I. Masser and M. Blakemore, pp. 78–102. Avon: Bath Press.

Shanholtz, V. O., C. J. Desai, N. Zhang, J. W. Kleene, and C. D. Metz (1990). "Hydrologic/Water Quality Modeling in a GIS Environment." ASAE Paper 90-3033, St. Joseph, Mich.: ASAE.

Skidmore, A. K. (1989). "A Comparison of Techniques for Calculating Gradient and Aspect from a Gridded Digital Elevation Model." *International Journal of Geographical Information Systems* 3(4), 323–34.

Smith, J. L., S. P. Prisley, and R. C. Weih (1991). "Considering the Effect of Spatial Data Variability on the Outcomes of Forest Management Decisions." *Proceedings of the GIS/LIS '91 Conference, Atlanta, Georgia*, pp. 286–92.

Srinivasan, R., and B. A. Engel (1991). "Effect of Slope Prediction Methods on Slope and Erosion Estimates." *Applied Engineering in Agriculture* 7(6), 779–83.

Taylor, J. R. (1982). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. Mill Valley, Calif.: University Science Books.

Theobald, D. M. (1989). "Accuracy and Bias Issues in Surface Representation." In *Accuracy of Spatial Databases*, edited by M. F. Goodchild and S. Gopal, pp. 99–106. London: Taylor & Francis.

USGS (1990). Digital Elevation Models: Data User's Guide 5. Reston, Va.: Earth Science Information Center, U.S. Geological Survey, 51 pp.

Whittle, P. (1950). "On Stationary Processes in the Plane." *Biometrika* 41, 434–49.

Wood, J. D., and P. F. Fisher (1993). "Assessing Interpolation Accuracy in Elevation Models." *IEEE Computer Graphics and Applications* 13(2), 48–56.

Zevenbergen, L. W., and C. R. Thorne (1987). "Quantitative Analysis of Land Surface Topography." *Earth Surface Processes and Landforms* 12, 47–56.