Recognition and Reduction of Systematic Error in Elevation and Derivative Surfaces from 7½-Minute DEMs

Daniel G. Brown and Thaddeus J. Bara

Abstract

The presence of systematic errors was observed in digital elevation data obtained from the USGS as 7½-minute quadrangle maps within a study area including the high relief environment of Glacier National Park, Montana. Digital surfaces of elevation and elevation derivatives, including slope angle and curvature, were examined for the presence of anisotropy through the use of calculated semivariograms and fractal dimensions. The existence of strongly anisotropic conditions analytically confirmed the presence of systematic errors in the dataset. The anisotropic conditions increased with the calculated derivative surfaces. Standard filtering procedures were applied to reduce the systematic errors. Alternative filters, prescribed after analysis of the semivariograms, reduced the magnitude of the anisotropy more than did a standard 3 by 3 low-pass filter.

Introduction

Systematic errors in digital data are the result of "some deterministic system which, if known, may be represented by some functional relationship" (Thapa and Bossler, 1992, p. 836). Systematic errors may be the most significant errors in a spatial or statistical analysis because they are not easily detected yet can introduce significant bias.

Two techniques, based on the assumption that short-range anisotropy in spatial data is indicative of error, are presented for revealing the presence and form of systematic error. Anisotropy is present when the general pattern of variation in one direction (eg., north to south) is different from the pattern of variation in another direction (eg., east to west). Webster (1985) discussed several approaches to analyzing anisotropy using semivariogram analysis. A simplified version of the conical projection discussed by Webster, involving the calculation of semivariance and fractal dimension in two orthogonal directions, is used here to assess the presence and degree of anisotropy in the original and derivative topographic surfaces.

DEM data produced by the U.S. Geological Survey are accuracy checked against the source products through standardized procedures, involving a comparison of elevation values from at least 20 randomly selected points within the digital dataset with the values at the same 20 locations on the source product (U.S.G.S., 1987). A minimum acceptable root-mean-square error (RMSE), or difference between DEM and source product, of seven metres has been established (U.S.G.S., 1987). This method of error calculation may be adequate under the assumption that errors are randomly distributed throughout the dataset.

The assumption of randomness in accuracy tests is confounded by the presence of systematic errors (Li, 1991). Aerial photograph scanning procedures used in the production of DEMs can result in "banding" or "striping" remnants, a form of systematic error process which commonly affects photogrammetric digital elevation data (U.S.G.S., 1987). The techniques presented in this paper are well suited to revealing and analyzing this type of systematic error.

Recognizing that filtering can reduce both systematic and random errors in a dataset, O'Callaghan and Mark (1984) applied a 3 by 3 low-pass filter to a DEM dataset subsequently used for drainage network extraction. While the 3 by 3 filter is a popular approach in dealing with error, it may be inadequate for removing all of the observed error. Suggested improvements include altering the filter window to a size which is more appropriate for the pattern of error as revealed in an analysis of anisotropy. Three low-pass filters (3 by 3, 3 by 5, and 3 by 7) were tested for their ability to reduce the systematic errors indicated by the semivariogram and fractal analysis. Suggestions for filtering USGS photogrammetric DEM datasets to reduce systematic error are provided.

Semivariogram and Fractal Concepts

Semivariance is the primary tool of modern geostatistics, a field of analysis developed from regionalized variable theory for the modeling of continuous, non-deterministic surfaces exhibiting spatial dependence. Geostatistics were initially applied to mining geology (Journel and Huijbregts, 1978; Clark, 1979), and were extended for analysis of the pattern and spatial structure of soilscapes (Burrough, 1983a; Burrough, 1983b; Webster, 1985), vegetation (Cohen *et al.*, 1990), and topographic surfaces (Mulla, 1988; Oliver *et al.*, 1989). Semivariances have been analyzed to suggest optimal cell sizes for modeling within a raster-based geographic information system (Brown *et al.*, in press).

Semivariance analysis describes a surface as the average squared difference of surface values which are a given distance, or lag, apart. Semivariance in a single dimension, $\gamma(h)$, is estimated by the expression

$$\gamma(h) = 1/2(N - h) \sum_{i=1}^{N-h} (z(i) - z(i + h))^2 \qquad (1)$$

where N is the number of points on the surface, $z(i)$ is the

D.G. Brown is with the Department of Geography, Michigan State University, Lansing, MI 48824-1115.

T.J. Bara is with the Department of Geography, University of North Carolina, Chapel Hill, NC 27599-3220.

Photogrammetric Engineering & Remote Sensing, Vol. 60, No. 2, February 1994, pp. 189–194.

0099-1112/94/6002–189$03.00/0
©1994 American Society for Photogrammetry and Remote Sensing

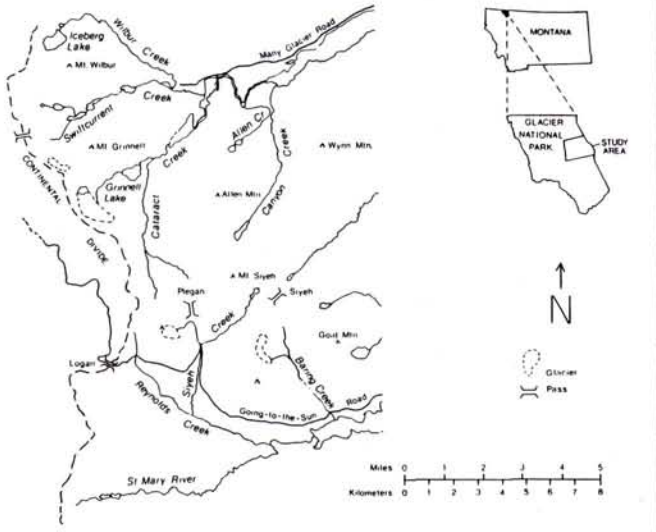

Figure 1. Study area map.

value of the surface at any point i, and $z(i+h)$ is the value of the surface h units from i. The shape of the semivariogram, a plot of semivariance against lag, reflects the underlying structure of spatial variability and is frequently modeled using spherical, linear, or exponential equations (Webster, 1985). Estimates of model parameters can be used as general descriptors of these underlying patterns (Mulla, 1988).

Statistical self-similarity across multiple spatial scales (the definition of a fractal which is used here) is present when semivariance increases logarithmically with lag (Burrough, 1983a). The fractal dimension of a surface can be calculated as a function of the slope of a log-log transform of the semivariogram by the equation

$$D = 3 - m/2 \quad (2)$$

where D is the fractal dimension and m is the slope of the log-log semivariogram. While topographic surfaces do not exhibit pure fractal behavior (Mark and Aronson, 1984), they often exhibit self-similarity across a limited range of scales. The calculations presented here are limited to variation within a range of 180 metres, within which fractal behavior was expected.

Polidori *et al.* (1991) used semivariograms and estimates of surface fractal dimensions for assessing systematic errors in DEMs. This paper demonstrates the logical extensions of their work by examining the effects of systematic error in the elevation data on errors in the derivative surfaces of elevation, slope angle, and slope curvature, and by suggesting ways of correcting for these errors.

Study Area

A portion of Glacier National Park (GNP), in northwestern Montana, was chosen as the study area for this analysis (Figure 1). Terrain is among the most important factors which regulate the hydrologic (Walsh *et al.*, 1990), ecologic (Brown, 1991; Bian and Walsh, 1992), and geomorphic (Butler and Walsh, 1990) processes active in GNP. The rugged landscape in the Park was shaped by Pleistocene mountain and valley glaciers. Hanging valleys, cirques, aretes, and horns are prominent features of the GNP landscape.

Elevations in the study area range from 1120 metres to 3000 metres. Rapid changes in elevation are common. Over small distances (i.e., less than 30 metres) slopes of 90 degrees are present (Butler and Walsh, 1990).

The study area encompassed parts of four 7½-minute quadrangles, including Logan Pass, Many Glacier, Lake Sherburne, and Rising Sun. Each of the DEMs was produced through the TRASTER procedure at the Rocky Mountain regional office of the USGS (U.S.G.S, 1987). Production involved scanning two stereoscopic aerial photographs and photogrammetrically plotting elevation values along parallel transects.

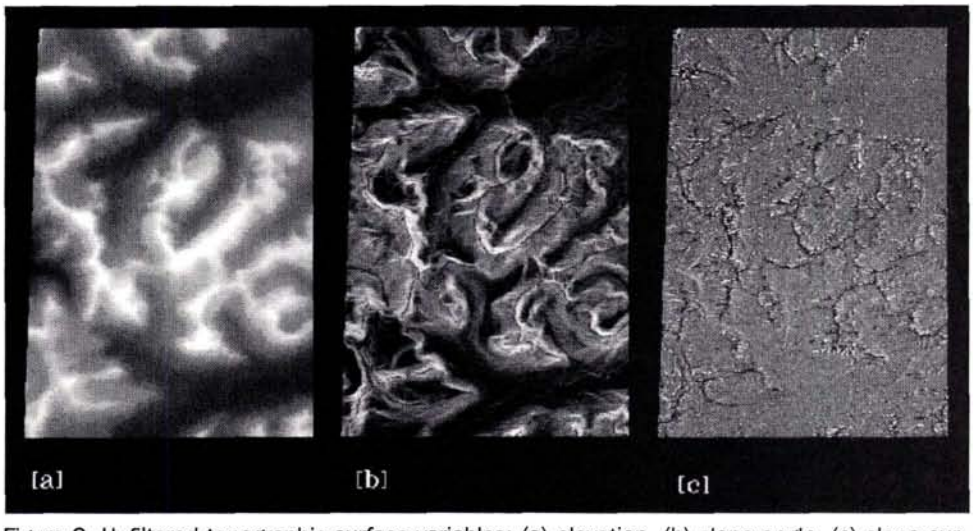

Figure 2. Unfiltered topographic surface variables: (a) elevation, (b) slope angle, (c) slope curvature.

The unfiltered elevation surface and its derivative slope angle and slope curvature surfaces for the study area are shown in Figure 2. Banding errors were particularly evident in the slope surface (especially in the southern quarter of the image) and the curvature surface, though reduction and reproduction necessary for publication conceal the errors somewhat.

Methodology

The DEMs were entered into ARC/INFO lattice format using DEMLATTICE, and converted to the ERDAS raster format using the ARC/INFO commands LATTICEGRID and SVFERDAS. All subsequent processing was carried out with software provided with the ERDAS system and with FORTRAN language programs, written using the ERDAS Programmer's Toolkit of subroutines (ERDAS, 1991).

Slope angles were calculated using the ERDAS TOPO module. A FORTRAN language program was written by the authors to generate a dimensionless, relative curvature surface from each elevation surface. Partial second-order derivatives were calculated in the east-west (x) and north-south (y) directions for each pixel on the elevation surface, using fourth-order finite difference calculus by the expressions

$$\text{partial } x = (1/12)(16z(i+1,j)+16z(i\text{-}1,j - 30z(i,j)-z(i-2,j) - z(i+2,j) \quad (3)$$

$$\text{partial } y = (1/12)(16z(i,j+1)+16z(j-1) - 30z(i,j) - z(i,j-2) - z(i,j+2) \quad (4)$$

where $z(i,j)$ is an elevation value at row and column location i,j. For peripheral cells lacking first- or second-order neighbors in one or more directions (e.g., edge pixels), second-order backward or forward finite difference calculus was used to calculate the partial derivatives (Hornbeck, 1975). The partial second derivatives were transformed to polar coordinates, where $r = \sqrt{}$ [partial(x)2 + partial(y)2] and θ = arctan [partial(y) / partial(x)]. The relative signs of the partial derivatives were used to assign θ to the appropriate quadrant. The transformed partials were projected onto a 45° to 135° axis of relative curvature. Finally, the values on this axis were scaled between 1 and 255, with 128 representing zero curvature (e.g., flat planar surface), 129 to 255 representing increasing concavity and 127 to 1 representing increasing convexity.

A FORTRAN language program was written to independently calculate semivariance as a function of lag distance in the two cardinal directions, north-south (N-S) and east-west (E-W). The semivariograms were calculated at 1-pixel lag intervals, out to a lag of 10 pixels (300 metres), for each of 12 data sets described below. The results of these calculations were exported to a graphing package for the production of the semivariograms.

Semivariograms were produced for the N-S and the E-W cases from elevation, slope angle, and slope curvature data for the detection of anisotropy. Fractal dimensions were calculated for each case using Equation 2 and the slope exponent of a best-fit logarithmic regression function through the semivariogram. Fractal dimensions were only calculated out to the sixth pixel lag (180 metres) to ensure a good regression fit (all but one of the fits had a Pearson R value of 0.99 or greater).

The elevation data were smoothed, using a standard, 3 by 3 low-pass filter (LPF) of the form

3 by 3 LPF

1	1	1
1	1	1
1	1	1

The smoothed elevation, as well as slope angle and slope curvature calculated from the smoothed elevation surface, were then tested for anisotropy through semivariogram analysis.

Two alternative smoothing kernels, a 3 by 5 and a 3 by 7 LPF, were subsequently applied to the original elevation surface. Both kernels were designed to provide additional smoothing in the N-S direction, the dominant orientation of the systematic errors. The forms of these filters are as follows:

3 by 5 LPF

1	1	1
1	1	1
1	1	1
1	1	1
1	1	1

3 by 7 LPF

1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1

The resultant smoothed elevation surfaces, as well as slope angle and slope curvature values calculated from those surfaces, were again examined through semivariogram and fractal analyses for the presence of anisotropic properties. Conclusions were drawn regarding the effectiveness of the three filters for reducing errors and the potential data loss resulting from the filtering process.

Results

Assuming that a dataset unaffected by systematic errors would be isotropic over short distances (i.e., exhibit similar spatially dependent variation in all directions), semivariograms were constructed to test for anisotropy in the unfiltered elevation and the slope angle and slope curvature datasets calculated from unfiltered elevation in the study area (Figures 3a, 4a, and 5a). If the datasets were isotropic, the shapes of the semivariograms constructed for the N-S and E-W dimensions would be similar. While this was somewhat true for the semivariograms of unfiltered elevation (Figure 3a), the slope angle or slope curvature calculated from unfiltered elevation exhibited significant anisotropy (Figures 4a and 5a, respectively). The cyclic form of the semivariogram for the curvature surface (Figure 5a) suggested that the systematic error was related to the banding artifact and that it was only present in the N-S transects.

Fractal dimensions calculated from the semivariograms (to a lag of 6 pixels) indicated the degree of disparity between N-S and E-W variance patterns in all three surfaces (Figure 6). The E-W transects exhibited lower fractal dimensions, indicating a higher degree of spatial dependence within 180 metres. The difference between the fractal dimensions increased progressively with the subsequent derivative surfaces of slope angle and slope curvature.

These analyses indicate that systematic error patterns present but not easily detected in the elevation data set were more noticeable after neighborhood operations (e.g., slope angle and slope curvature calculations) had been performed. Although it is difficult to quantify the degree of error magnification, the finding that systematic errors can be increased

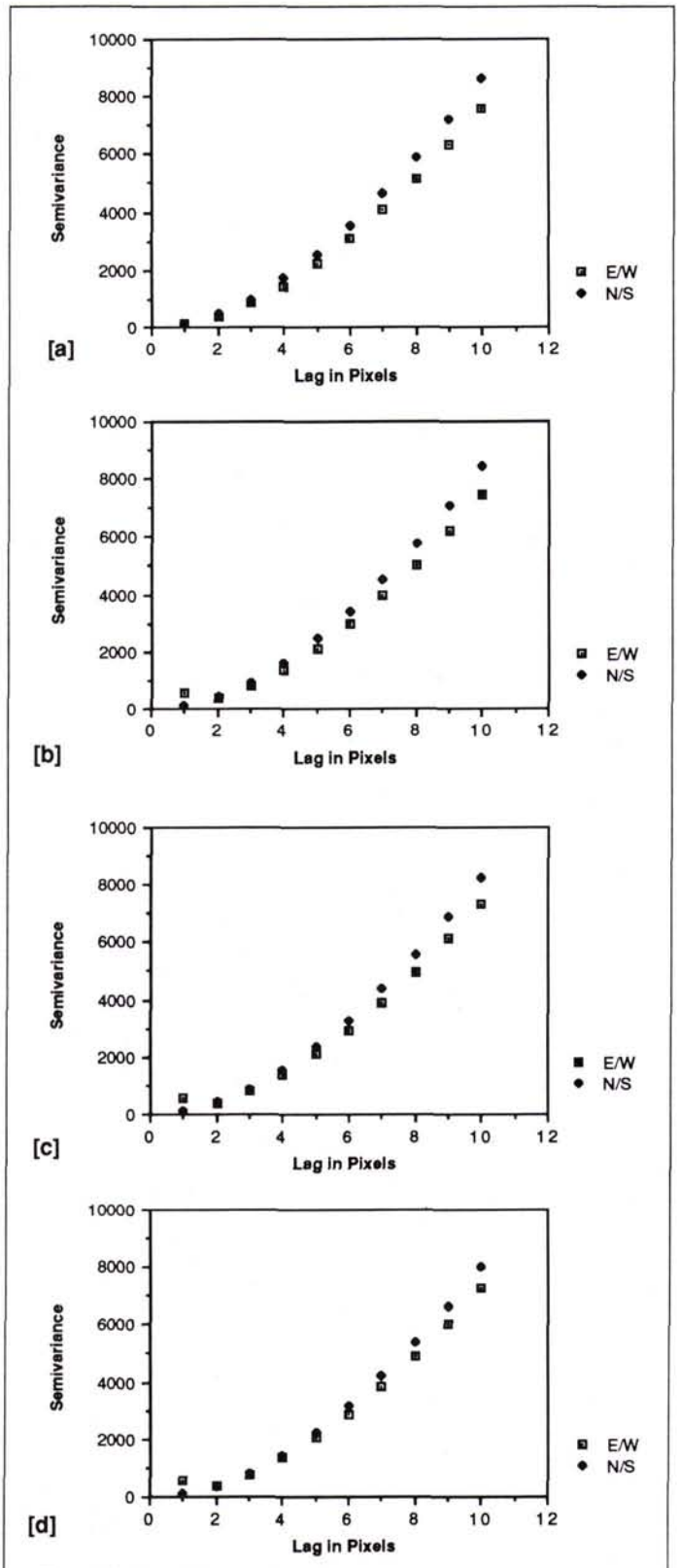

Figure 3. East/West and North/South semivariograms for elevation: (a) unfiltered, (b) 3 by 3 LPF, (c) 3 by 5 LPF, (d) 3 by 7 LPF.

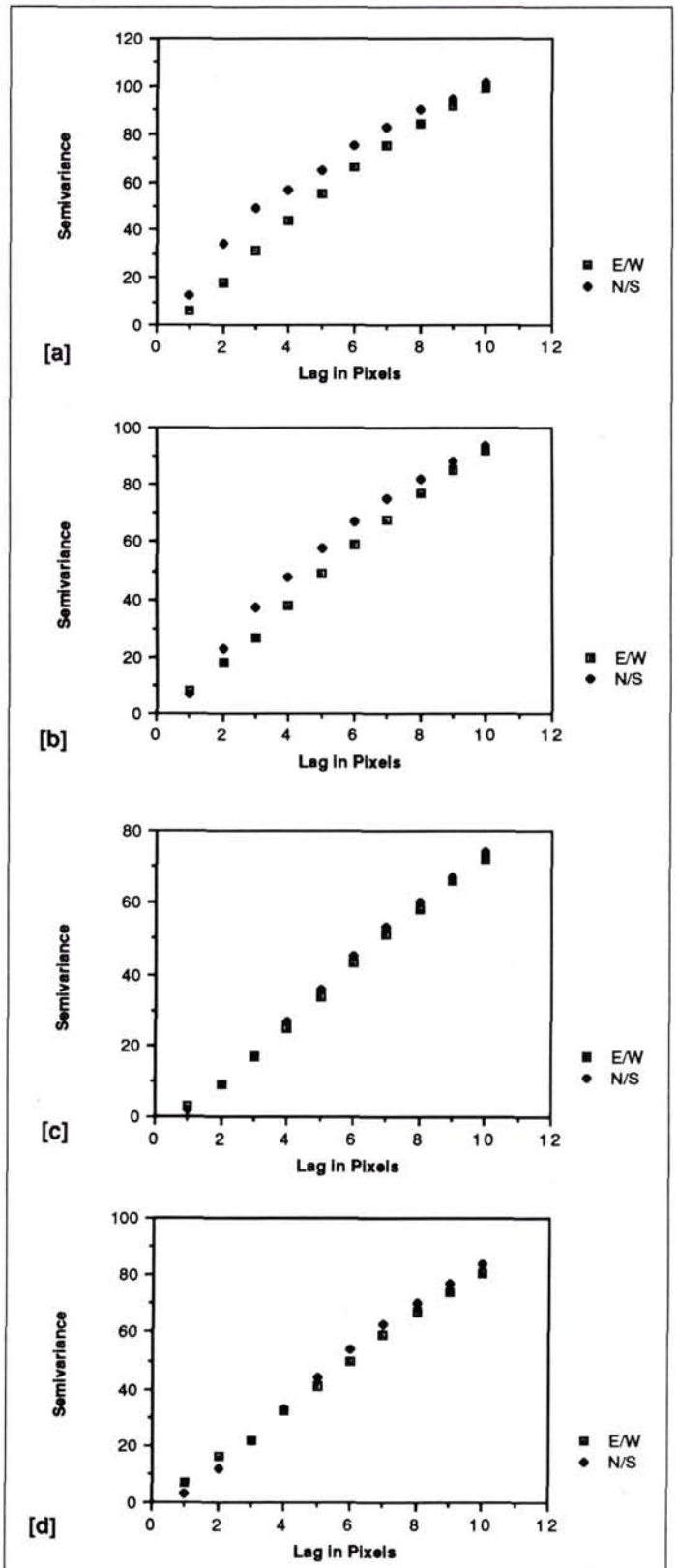

Figure 4. East/West and North/South semivariograms for slope angle: (a) unfiltered, (b) 3 by 3 LPF, (c) 3 by 5 LPF, (d) 3 by 7 LPF.

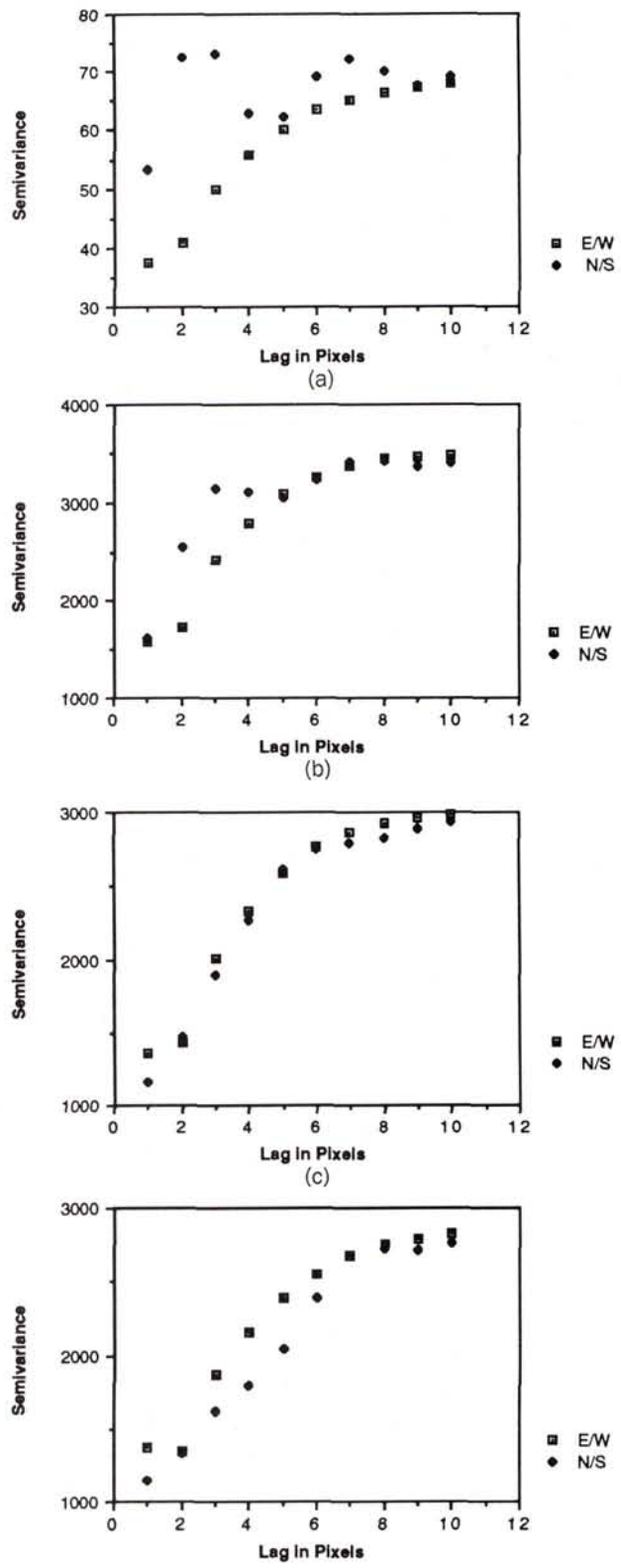

Figure 5. East/West and North/South semivariograms for slope curvature: (a) unfiltered, (b) 3 by 3 LPF, (c) 3 by 5 LPF, (d) 3 by 7 LPF.

through the calculation of derivative surfaces concurs with the conclusions of others (Walsh *et al.*, 1987; Lanter and Veregin, 1992).

Filtering to Reduce Error

Low-pass filters (LPFs) are suggested as a device for reducing the systematic errors present in the DEM dataset and derivatives. The three filters (3 by 3, 3 by 5, and 3 by 7 LPFs) were applied to the elevation data and resulted in the semivariogram patterns shown in Figures 3b through 3d. The forms of the semivariograms for the derivative surfaces of slope angle and slope curvature are given in Figures 4b through 4d and Figures 5b through 5d.

Although the N-S and E-W elevation semivariograms tended to converge somewhat with increasing N-S filter size, the effect is more clearly evident in the slope angle and slope curvature semivariogram forms. The N-S semivariograms of slope and curvature, calculated from the 3 by 3 LPF elevation data (Figures 4b and 5b), deviated from the E-W pattern less than the patterns calculated from the unfiltered elevation. The 3 by 5 LPF had the effect of nearly equalizing the N-S and E-W patterns (Figures 4c and 5c). Finally, the 3 by 7 filter appeared to have over-compensated for the anisotropic pattern, affecting N-S elevation variation to the point that both slope and curvature were excessively smoothed in the N-S dimension (Figures 4d and 5d).

The fractal dimensions for elevation, slope angle, and slope curvature from the 3 by 3, 3 by 5, and 3 by 7 LPF datasets confirmed these observations (Figure 6). The N-S fractal dimension values were higher than the E-W values in the 3 by 3 LPF elevation data and associated derivatives, though the differences in the fractal dimensions were reduced from the values in the unfiltered data. The 3 by 5 LPF fractal dimensions in both orientations were nearly equivalent. Differences in the 3 by 7 LPF fractal dimensions for the elevation and slope surfaces were reduced with the N-S transects having lower values. The 3 by 7 LPF curvature semivariogram exhibited a poor logarithmic fit and was not, therefore, as reliable as the other values.

Differences between Filtered and Unfiltered Data

While filtering improves the spatial structure of the datasets and reduces systematic error, filtering may have had the undesirable effect of altering elevation values which were actually correct. The potential amount of information loss which may have occurred as the result of filtering was as-

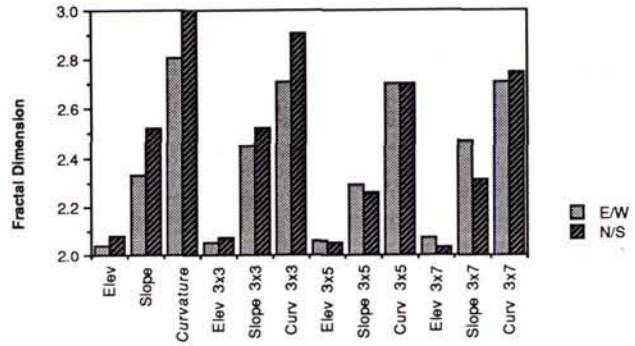

Figure 6. East/West and North/South fractal dimensions for each topographic variable and each filter.

TABLE 1. CENTRAL TENDENCY MEASURES OF THE ABSOLUTE DIFFERENCE DISTRIBUTION BETWEEN FILTERED AND UNFILTERED ELEVATION VALUES.

	Mean	Median (in metres)	Mode
3 by 3 LPF	2.9	1	1
3 by 5 LPF	4.7	2	1
3 by 7 LPF	6.4	3	1

sessed by comparing filtered and unfiltered elevation surfaces. Three absolute difference surfaces were constructed by calculating the absolute value of the difference between the filtered (3 by 3, 3 by 5, and 3 by 7 LPFs) and unfiltered elevation values at each pixel. Central tendency measures (mean, median, and mode) were calculated for the absolute difference surfaces (Table 1).

As expected, the central tendency of differences between filtered and unfiltered elevation values increased with increasing filter size. While the filtering process increased uncertainty in the elevation values, it is difficult to formally assess the accuracy of the filtered elevation surfaces without comparing them with the control points used by the USGS for accuracy assessment. It is useful, nonetheless, to note that the changes in elevation values with filtering appear to be minor, with the mode of each absolute difference surface being 1 metre and the mean difference being less than 6.5 metres for each filtered surface.

Conclusions

The local-scale anisotropic pattern within USGS DEM products studied in this paper are consistent with patterns resulting from banding, a common by-product of DEMs generated through photogrammetric procedures. These errors are seen to increase in severity with the calculation of derivative surfaces and have the potential to result in significant distortions and biases in analyses which use these products. Consequently, the acknowledgement of the presence of these errors and methods to correct them are of interest to all users of digital elevation products generated by photogrammetric techniques.

This paper illustrated how semivariance and fractal analysis over short ranges can be used to

- detect the presence and structure of the systematic errors of DEM products, and
- provide a quantitative basis for applying corrections to the surface to mitigate the severity of these errors.

Smoothing kernels were generated and applied using standard features of a commercially available image processing/GIS system. These kernels were shown to have the desirable effect of decreasing the observed anisotropy in the elevation and derivative surfaces, while resulting in only small absolute differences between the original and smoothed elevation surfaces. For the DEMs evaluated in this study, a 3 by 5 low-pass filter (LPF) oriented in the N-S direction resulted in comparable semivariogram forms and fractal dimensions in the N-S and E-W directions, with a mean absolute vertical deviation of 4.7 metres between the unfiltered and filtered elevation surfaces.

Acknowledgments

This work was financially supported, in part, by the Global Change Research Program of the Earth Science and Applications Division of NASA (Grant Number NGT-30068). In addition, computing facilities for this research were provided by Dr. Stephen J. Walsh at the University of North Carolina.

References

Bian, L., and S. J. Walsh, 1993. Scale dependencies of vegetation and topography in a mountainous environment, *The Professional Geographer*, Vol. 45, No. 1, pp. 1–11.

Brown, D. G., 1991. Topoclimatic models of an alpine environment using digital elevation models within a GIS, *Proceedings GIS/LIS '91 Conference*, pp. 835–844.

Brown, D. G., L. Bian, and S. J. Walsh, in press. Response of a Distrubuted Watershed Erosion Model to Variations in Input Data Aggregation Levels, *Computers and Geosciences*.

Butler, D.R., and S.J. Walsh, 1990. Topographic, structural and lithologic influences of snow avalanche path location, eastern Glacier National Park, Montana, *Annals of the Association of American Geographers*, Vol. 80, No. 3, pp. 362–378.

Burrough, P. A., 1983a. Multiscale sources of spatial variation in soil: I. The application of fractal concepts to nested levels to soil variation, *Journal of Soil Science*, Vol. 34, pp. 577–597.

———, 1983b. Multiscale sources of spatial variation in soil: II. A non-Brownian fractal model and its application in soil survey, *Journal of Soil Science*, Vol. 34, pp. 599–620.

Clark, I., 1979. *Practical Geostatistics*, Applied Science, London.

Cohen, W. B., T. A. Spies, and G. A. Bradshaw, 1990. Semivariograms of digital imagery for analysis of conifer canopy structure, *Remote Sensing of Environment*, Vol. 34, pp. 167–178.

ERDAS, 1991. *ERDAS Field Guide*, second edition. ERDAS, Inc., Atlanta.

Hornbeck, R. W., 1975. *Numerical Methods*, Prentice Hall, Englewood Cliffs, N.J..

Journel, A. J., and C. J. Huijbregts, 1978. *Mining Geostatistics*, Academic Press, London.

Lanter, D. P., and H. Veregin, 1992. A research paradigm for propagating error in layer-based GIS, *Photogrammetric Engineering & Remote Sensing*, Vol. 58, No. 6, pp. 825–833.

Li, Z., 1991. Effects of check points on the reliability of DTM accuracy estimates obtained from experimental tests, *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 10, pp. 1333–1340.

Mark, D. M., and P. B. Aronson, 1984. Scale-dependent fractal dimensions of topographic surfaces: an empirical investigation with application in geomorphology and computer mapping, *Mathematical Geology*, Vol. 16, pp. 671–683.

Mulla, D. J., 1988. Using geostatistics and spectral analysis to study spatial patterns in the topography of southeastern Washington State, USA, *Earth Surface Processes and Landforms*, Vol. 13, pp. 389–405.

O'Callaghan, E. M., and D. M. Mark, 1984. The extraction of drainage networks from digital elevation data, *Computer Vision Graphics and Image Processing*, Vol. 28, pp. 323–344.

Oliver, M. A., and R. Webster, 1989. Geostatistics in physical geography. Part II: applications, *Transactions of the Institute of British Geographers*, Vol. 14, pp. 270–286.

Polidori, L., J. Chorowicz, and R. Guillande, 1991. Description of terrain as a fractal surface, and application to digital elevation model quality assessment, *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 10, pp. 1329–1332.

Thapa, K., and J. Bossler, 1992. Accuracy of spatial data used in geographic information systems, *Photogrammetric Engineering & Remote Sensing*, Vol. 58, No. 6, pp. 835–841.

U.S.G.S., 1987. *Digital Elevation Models, Users Guide 5*. U.S. Geological Survey, Reston, Virginia.

Walsh, S. J., D. R. Lightfoot, and D. R. Butler, 1987. Recognition and assessment of error in geographic information systems, *Photogrammetric Engineering & Remote Sensing*, Vol. 53, No. 10, pp. 1423–1430.

Walsh, S. J., J. W. Cooper, I. E. Von Essen, and K. R. Gallager, 1990. Image enhancement of Thematic Mapper data and GIS data integration for evaluation of resource characteristics, *Photogrammetric Engineering & Remote Sensing*, Vol. 56, No. 8, pp. 1135–1141.

Webster, R., 1985. Quantitative spatial analysis of soil in the field. *Advances in Soil Science*, Vol. 3, pp. 1–70.

(Received 6 July 1992; accepted 26 January 1993)