METHODS AND APPLICATIONS IN SURFACE DEPRESSION ANALYSIS

Susan K. Jenson*
TGS Technology, Inc.
Sioux Falls, South Dakota 57198

and

Charles M. Trautwein
U.S. Geological Survey
EROS Data Center
Sioux Falls, South Dakota 57198

ABSTRACT

Gridded surface data sets are often incorporated into digital data bases, but extracting information from the data sets requires specialized raster processing techniques different from those historically used on remotely sensed and thematic data. Frequently, the information desired of a gridded surface is directly related to the topologic peaks and pits of the surface. A method for isolating these peaks and pits has been developed, and two examples of its application are presented.

The perimeter of a pit feature is the highest-valued closed contour surrounding a minimum level. The method devised for finding all such contours is designed to operate on large raster surfaces. If the data are first inversely mapped, this algorithm will find surface peaks rather than pits.

In one example the depressions, or pits, expressed in Digital Elevation Model data, are hydrologically significant potholes. Measurement of their storage capacity is the objective. The potholes are found and labelled as polygons; their watershed boundaries are found and attributes are computed.

In the other example, geochemical surfaces, which were interpolated from chemical analyses of irregularly distributed stream sediment samples, were analyzed to determine the magnitude, morphology, and areal extent of peaks (geochemical anomalies).

---

* Work performed under U.S. Geological Survey contract no. 14-08-0001-20129.

Publication authorized by the Director, U.S. Geological Survey. Any use of trade names and trademarks in this publication is for descriptive purposes only and does not constitute endorsement by the U.S. Geological Survey.

RATIONALE

Gridded surface data sets are critical components in many digital spatial databases. For example, Digital Elevation Model (DEM) data may be used to derive hydrologic information (Jenson, 1984), and gridded geochemical surfaces may be used to delineate areas of anomalous concentrations of a chemical element (Dwyer and others, 1984). However, while gridded surfaces are valuable datasets, they require analytical tools that recognize their special characteristics. While discontinuities may be present, gridded surfaces are discrete representations of primarily continuous data. The original control data that are used to generate a surface may be contour lines or control points, but the algorithms that compute the surfaces assume a continuous model. This underlying assumption distinguishes gridded surfaces from thematic spatial information such as digitized lithological units, and from remotely sensed data where manmade and natural discontinuities are frequent.

The information extracted from gridded surfaces is typically of a continuous nature, reflecting the data's origins and assumptions. For instance, slope and aspect information is commonly computed from DEM's, and directional derivatives are computed for geophysical surfaces. For visual information extraction, surface data are often represented using contour maps and mesh diagrams that aid interpretation of surface highs and lows.

The analytical tool presented here finds the topologic peak and pit polygonal features of a gridded surface. It is a specialized contouring process because the perimeters that define the peak and pit features are the lowest possible and highest possible closed contours, respectively, of the surface. Since these perimeters may occur at any data value in the data range, it is not possible to find these contours with standard algorithms without using an unreasonably small contour interval. This procedure was developed for hydrologic studies with DEM's, as in the DEM example presented later in the paper; however, it has utility for other types of surface data as well, as illustrated in the geochemical example.

An algorithm developed by Chan (1985) locates "lakes" in DEM data by locating a cell which may be within a depression and "growing" the lake with a stack-oriented algorithm. This approach requires an unacceptably large amount of computer memory to be allocated when depressions are large, as in some of the data presented here.

THE PROCEDURE

Since finding peak areas is the opposite process to finding pit areas, the peak- and pit-finding procedure was divided into two steps, both of which use the same computer programs. To find peak areas, the data are first inversely mapped. Either step or both steps may be selected for a given application. For purposes of describing the procedure, it will be assumed to be producing pit areas.

The objective of the procedure is to produce an output raster surface identical to the input raster surface, but with the cells contained in depressions raised to the lowest value on the rim of the depression. Therefore, each cell in the output image will have at least one monotonically decreasing path of cells leading to an edge of the data set. A path is composed of cells that are adjacent horizontally, vertically, or diagonally in the raster (eight-way connectedness) and that meet the steadily decreasing value criteria. If the input surface is subtracted from the output surface, each cell's resulting value is equal to its depth in a depression in the units of the input surface.

In order to accommodate large surfaces, the program was designed to operate in two modes. In the first mode, the surface is processed by finding and filling depressions wholely contained in 100-line by 100-sample blocks. In the second mode, the entire surface is processed iteratively in a circular buffer. Use of the first mode is memory intensive and the second is input and output intensive. A data set is first processed by the program in the first mode, thereby filling all depressions that do not intersect with cells with line or sample coordinates that are evenly divisible by 100. This intermediate data set is then processed by the program in the second mode to fill the remaining depressions. Processing in the second mode requires that only four lines of data be resident at any one time; therefore, large images can be processed. It is possible to further optimize the process for a given data set by varying the number of lines and samples processed in the first mode and to repeat the first mode with blocks staggered to overlie the join lines of the previous first mode pass. These modifications allow more of the depressions to be filled in the more efficient first mode. The procedure by which the first mode processes a block and the second mode processes the entire surface is the same for both modes and is as follows:

1. Mark all cells on the data set edges as having a path to the edge.

2. Mark all cells that are adjacent to marked cells and are equal or greater in value. Repeat this step until all possible cells have been marked.

3. Find and label all eight-way connected polygons of unmarked cells such that each polygon has maximum possible area. If no polygons are found, end the procedure.

4. For each polygon, record the value of the marked cell of lowest value that is adjacent to the polygon (threshold value).

5. For each polygon, for each cell in the polygon, if the cell has a value that is less than the polygon's threshold value, then raise the cell's value to the threshold value.

6. Repeat from step 2.

GEOCHEMICAL APPLICATION

An application to the detection and spatial characterization of geochemical anomalies that has been investigated demonstrates the utility of automated depression analysis techniques in the analysis of complex geochemical terrains. Geochemical anomalies, commonly defined by unusually high local concentrations of major, minor, and trace elements in rocks, sediments, soils, waters, and atmospheric and biologic materials, are important features in studies related to mineral and energy resource exploration and environmental monitoring. These anomalies are usually detected by establishing a threshold concentration that marks the lower bound of the anomalous concentration range for each element in each type of material. The threshold value is used to sort the geochemical data into background and anomalous sample populations, which then may be plotted on a map for comparison with other data. For certain types of materials and terrains in which background concentrations for selected elements are relatively uniform, this approach is satisfactory. However, in geochemical terrains where background values are variable across the region studied, this approach is commonly modified by removing regional trends prior to the selection of an appropriate threshold value.

Trend-surface analysis is frequently used to mathematically model regional variations in geochemical data sets. In this technique, first-, second-, and higher-order equations are used to describe regional trends in terms of the data set's best least-squares fit to planar, parabolic, and higher order nonplanar surfaces. The resultant regional model is subtracted from the original data leaving residual concentrations that represent local variations, above and below, the regional trend. Positive variations are then statistically evaluated to establish a threshold. This procedure works well in areas where regional controls, and their consequent effects, are known; however, in most areas trend-surface analysis only provides an approximation of an unknown function with an arbitrary, best-fit function.

Because many types of geochemical data are cartographically represented as contour maps (with contour intervals equated to chemical concentration ranges) and geochemical anomalies are topologically analogous to localized peaks on a topographic map, automated depression analysis techniques were applied to a rasterized geochemical data set in an effort to more objectively define anomalies based on their morphology.

A geochemical data set was studied that consisted of 2,639 analyses of copper concentration in the heavy mineral fraction of stream sediment samples distributed throughout the Butte 1° x 2° Quadrangle, Montana. The analyses, which were referenced by latitudes and longitudes of the sample collection sites, were rasterized using a minimum curvature interpolation and gridding algorithm (Briggs, 1977). The resultant grid consisted of a 559- by 775-cell array of 200-meter by 200-meter (ground-equivalent size) grid cells cast in a Transverse Mercator map projection. Interpolated copper concentration values in the array were in the range from 0 to 65,684 ppm (parts per million) copper with an arithmetic mean of 161.17 ppm. Figures 1 and 2 show the distribution of original sample sites within the quadrangle and a grey-level representation of the interpolated concentration surface.

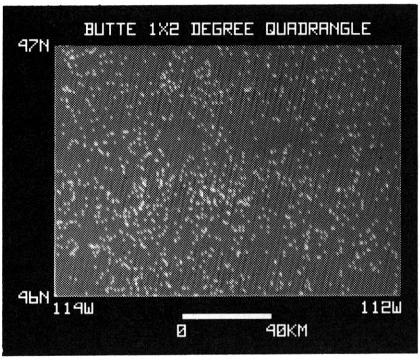

Figure 1.--Distribution of geochemical sample sites.

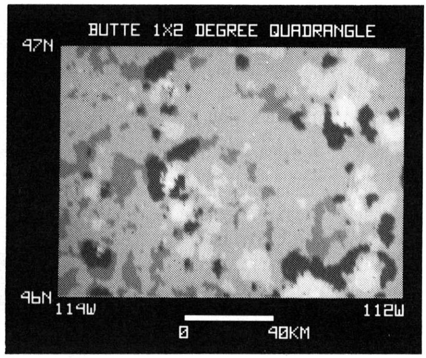

Figure 2.--Gray-level map of copper surface; brighter tones represent higher concentration intervals.

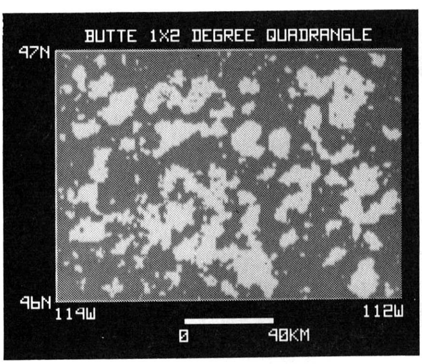

Figure 3.--Topologically defined copper anomalies.

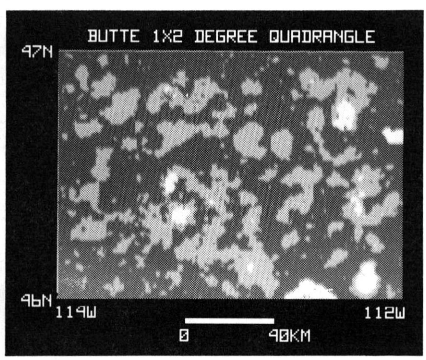

Figure 4.--Comparison of topologically defined copper anomalies (gray) and copper anomalies defined by a 1,000 ppm threshold (white).

A mapping function was used to topographically invert the interpolated range of values. The product of this operation was then subjected to the depression analysis algorithm described earlier. Anomalies, in their inverted form, are morphologically described through this algorithm as closed depressions. The depressions found in the inverted data for the Butte quadrangle are shown in gray in figure 3. In figure 4, these same depressions are shown in gray again, and superimposed in white are the areas that are above a 1,000 ppm threshhold. The white areas are the only areas that are identified by a traditional single-threshhold approach. This comparison demonstrates the utility of the morphologic approach in areas such as this where regional variations are extreme and a single threshold value is insufficient for detecting anomalies in different parts of the geochemical terrain. While the morphologic approach identifies many more potentially anomalous areas, more analysis is required to relate the ppm values in the area to the area's background material types. An additional advantage of this approach is that it does not require generation of a separate, often arbitrary, model of the regional trend as in cases where trend-surface analysis is performed.

A visual comparison of peaks and the control points that are within them or nearby them is beneficial in that each peak's reliability can be evaluated. If many control points appear to be defining a peak, the analyst may feel more confident in categorizing that peak as anomalously high. However, if the control points are few or badly distributed, the peak may be categorized as an overshoot in the surface-generation process.

DEM HYDROLOGY APPLICATION

The National Mapping Division and Water Resources Division of the U.S. Geological Survey cooperated with the Bureau of Reclamation in 1985 and 1986 to objectively quantify and to incorporate the contributing and noncontributing factors of pothole terrain in a probable maximum flood estimate for the James River Basin above the dam at Jamestown, North Dakota. The hydrology of the area has been difficult to study due to flat slopes, the complex nested drainage of the potholes, and a poorly defined drainage network.

DEM's were used to derive hydrologic characteristics that were incorporated in rainfall runoff models. DEM's were made for five test sites in the Basin. Each test site covered approximately 10 square miles with a 50- by 50-foot grid-cell size. The largest DEM was 505 lines by 394 samples.

For each test site, the surface depression procedure was the beginning step for the DEM analysis. Once the surface depressions were identified, they were given unique identifying labels and their volumes were calculated. A subset of depressions were selected for the modeling process based on a minimum volume criteria. Some depressions that did not meet the volume criteria were still modelled because they were spatially necessary to complete drainage linkages.

A second processing step then found the watershed boundaries for these selected depressions. A previous watershed program (Jenson, 1984) had to deal with real and artificial depressions in the paths of drainages by running iteratively and using thresholds to "jump" out of holes. By taking advantage of the depression map of the surface, however, the watershed program could be modified to run in two passes. The surface processed by the watershed program was the surface with all depressions filled except those that were selected for the modeling process. A shaded-relief representation for the DEM of one of the test sites is shown in figure 5. The corresponding selected potholes and watershed boundaries are shown in figure 6.

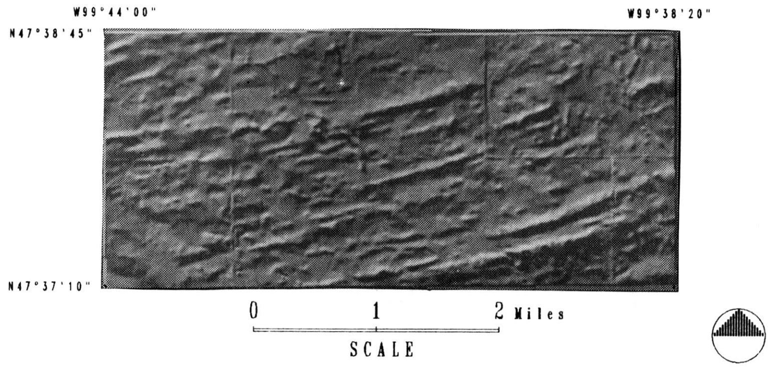

Figure 5.--Shaded-relief representation of digital elevation model data for one of the James River Basin test sites.

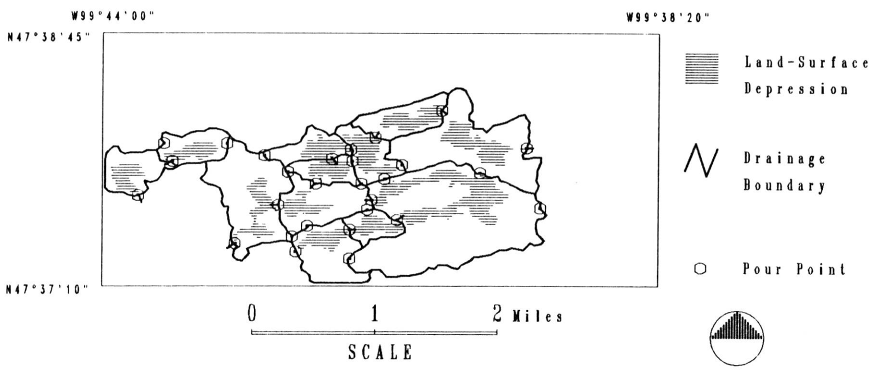

Figure 6.--Selected potholes, watersheds, and pour points for the test site in figure 5.

CONCLUSIONS

This depression-finding procedure has been shown to be practical and useful in the analysis of geochemical and DEM surface data sets. For inversely-mapped geochemical surfaces, depression analysis indicates areas of anomalously high chemical concentrations and bypasses the need for trend surface analysis. The hydrologic analysis of DEM surfaces benefits from depression identification because depressions may be hydrologically significant themselves, such as potholes, and the removal of unwanted depressions simplifies the automated finding of watershed boundaries.

REFERENCES

Briggs, I. C., 1977, Machine contouring using minimum curvature, Geophysics, vol. 39, no. 1, p. 39-48.

Chan, K. K. L., 1985, Locating "lakes" on digital terrain model; Proceedings, 1985 ACSM-ASPRS Fall Convention, p. 68-77.

Dwyer, J. L., Fosnight, E. A., and Hastings, D. A., 1984, Development and implementation of a digital geologic database for petroleum exploration in the Vernal Quadrangle, Utah-Colorado, U.S.A.; Proceedings, International Symposium on Remote Sensing of Environment, p. 461-475.

Jenson, S. K., 1984, Automated derivation of hydrologic basin characteristics from digital elevation model data; Proceedings, Auto-Carto 7, p. 301-310.