Contour Mapping and SURFACE II

JOHN C. DAVIS

A contour map is a diagram used to represent the form of the earth's surface in an area. Topography within the map area is represented by contours or lines of equal elevation drawn as though the landscape were successively flooded to greater depths and tracings made of the resulting shorelines. Contour maps are among the primary working tools of geologists, geographers, civil engineers, and others who are concerned with the configuration of the land.

The same graphical convention can be used to represent the forms of other types of surfaces, such as the top of a buried geologic stratum or the spatial variation of barometric pressure. The convention also can be extended to nongeographical coordinates to yield displays of one variable in terms of two others. Although it is more appropriate to refer to the form lines on such diagrams as isarithms, the term contour line is commonly used. Figure 1, for example, shows a contour map of the potential energy surface of a hydrogen molecule reacting with a hydrogen atom (*1*).

Often, contour maps must be produced from observations of the surface made at irregularly spaced locations called control points. It is necessary to interpolate between these control points to intermediate locations where contour lines are to be drawn. Performing the interpolations and drawing a contour map is an onerous task, one that geologists quickly turned over to computers and digital plotters when these became widely available in the early 1960s (*2*). In a market created primarily by oil companies, computer programs for contour mapping rapidly evolved into large, sophisticated, and very expensive software packages for the mainframe computers of the time.

The situation is quite different now. The proliferation of personal computers has led to development of relatively inexpensive programs for producing contour maps. A recent survey (*3*) identified over 30 contouring programs now available for IBM-compatible microcomputers alone. Although contouring programs may sell for prices that range from less than $100 to more than $50,000, they all embody the same basic principles. A higher price buys more elegant graphics, easier file manipulation, and (usually) greater efficiency.

Principles of Operation

The initial step in drawing a contour map is to construct a mesh or network of straight lines, either between pairs of control points with known values or between pairs of points whose values have been estimated. The locations where contours should cross the network are then found by interpolation. Finally, contour lines are drawn by connecting all intersection points that have the same values.

A basic difference between programs is in the nature of the mesh of lines across the map; the mesh may be either regular or irregular. An irregular mesh is composed of triangles, much like a surveyor's network, with each vertex defined by the location of a known control point. The form of the surface within a triangular area is modeled by a spatial function such as a bicubic spline that is calculated so that the surfaces within adjacent triangular patches blend smoothly without discontinuities (*4*). Once the coefficients for all surface patches have been calculated, contour lines can be traced across them at any desired resolution. This type of contouring is sometimes called deterministic, because the form of the surface is uniquely determined for any set of control points. The surface will pass exactly through every control point, will have minimum curvature, and will have continuous first and second derivatives (*5*).

The form of a surface contoured by this method will depend not only on the values at control points, but also on the spatial arrangement of these points. Merely connecting the control points in a different order will change the triangular mesh and result in a different contoured surface. Soft-

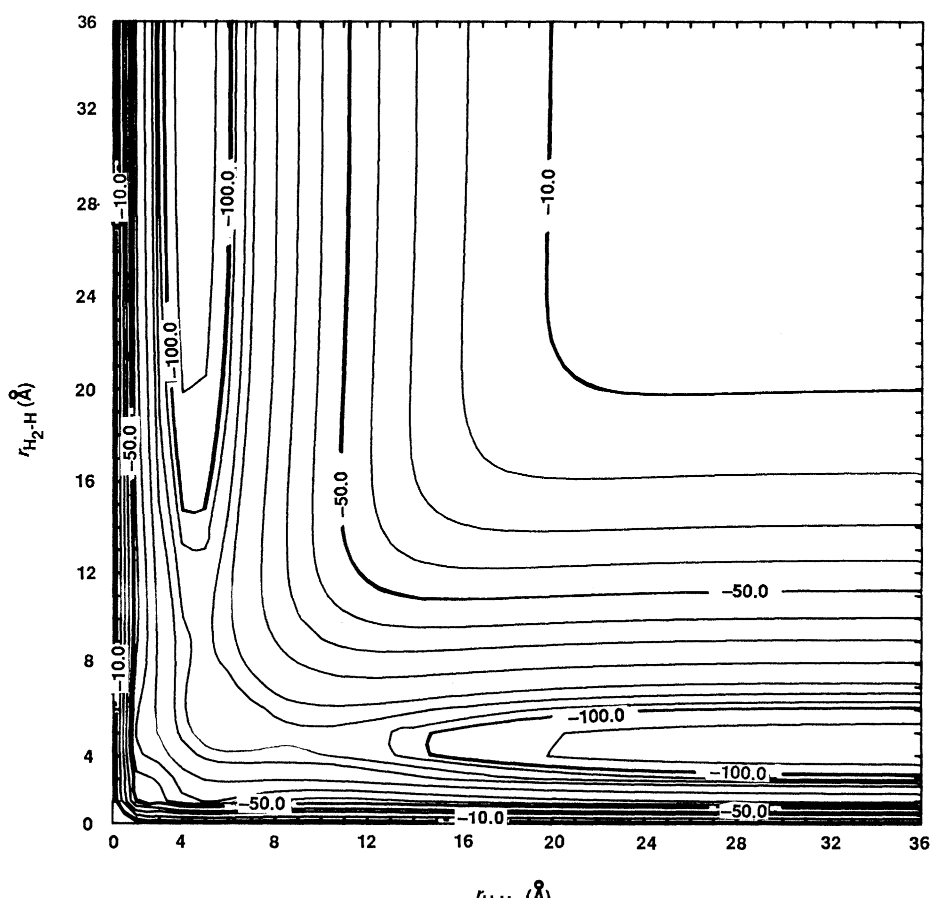

Fig. 1. Potential energy surface of a hydrogen molecule reacting with a hydrogen atom. The contour interval is 10 kcal/mol (*1*).

Advanced Projects Section, Kansas Geological Survey, Lawrence, KS 66046.

ware designers usually attempt to avoid this problem by defining a unique mesh based on Delaunay triangles, but if the data (or part of it) are collected at equally spaced locations, then ambiguities remain (*6, 7*). Another problem may arise in highly uneven arrangements of the data locations. If two control points are close together, they may form the base of a very elongate triangle. A small difference in values at these points may create a thin, sliverlike area of steep slope, where contour lines bend abruptly. The resulting artifices have an obvious relation to the placement of control points.

However, there is a more practical reason why deterministic algorithms have not become the procedure of choice for computer contouring. In many applications it is insufficient to create a representation of a single surface; rather, a comparison between two (or more) surfaces is required. For example, to calculate the volume of oil in a subsurface reservoir, the thickness of the oil-bearing interval must be estimated throughout its extent. This can be done by subtracting a map of the lower oil/water contact from a structural map of the upper surface of the reservoir unit. If lateral variations in porosity in the reservoir are also mapped, the map of reservoir thickness can be multiplied by the map of porosity to give a map of thickness of oil in place. Numerical integration will yield the total volume of oil. In other applications, the sum, difference, product, or logical combination of two maps may be required. These cannot be found easily from contour maps based on an irregular triangular mesh unless all maps have control points at exactly the same locations. This places a severe limitation on the usefulness of contouring programs that use the triangulation approach.

The alternative method for computer contouring uses a regular (usually square) array of intermediate points to guide the placement of contour lines. Since in general the locations of control points will not coincide with these regularly spaced points, their values must be estimated from the observations. This step usually is referred to as gridding, and the regularly spaced estimates of the surface are called grid nodes. Once these grid nodes have been found, contour lines are drawn by procedures similar to those used in deterministic algorithms, except that the surface patches are regular squares rather than irregular triangles. The final placement of contour lines depends upon estimated values of the surface at the regularly spaced grid nodes and not on the known values of the surface at the control points.

Although a gridding procedure is more complicated than the deterministic method, it offers several advantages. Since the intermediate grid is completely uniform, aberrant patterns of contour lines caused by poor spacing of control points do not occur. Also, because any mapped area can be covered by a grid with any desired spacing and orientation, two surfaces can be compared easily regardless of the number and placement of the control points in the area. Finally, gridding-type algorithms were until recently computationally faster than triangulation algorithms because finding the Delaunay triangulation of a random field of points was a lengthy trial-and-error process. Computer run times were dramatically shortened following development of the Green-Sibson triangulation algorithm (*6*), but even so contoured maps must be gridded deterministically before two maps can be compared.

At present, most commercial contouring programs designed for mainframe or minicomputer systems use an intermediate gridding step. Both gridding and triangulation-type programs are used widely on personal computers. Because triangulation programs do not store an intermediate grid, they require less memory than do gridding algorithms, an important consideration on the smaller machines.

SURFACE II

Because commercial contouring programs are expensive to purchase and time-consuming to install, side-by-side comparisons of their features and performance are rarely undertaken, although the Denver Geophysical Society recently sponsored a competition in which vendors demonstrated 13 mainframe and 10 personal-computer contouring programs (*8*). This review examines one program that has features that are typical of gridding-type contouring programs written for mainframes and minicomputers. SURFACE II was created by Robert Sampson in 1973 for the Kansas Geological Survey and has been distributed widely, especially to universities and research institutions. The list of registered users includes over 350 computer centers worldwide (although sales of the user's manual suggest a larger user group). The program has been popular in part because the Fortran-77 source code is provided and because nonprofit organizations can obtain the program for a nominal fee. SURFACE II was designed for research on problems of contouring as well as for the routine production of contour maps so that the program has many statistical features and options not available in other programs (*9*). This has contributed to its use as a teaching and research tool, especially by departments of geography and cartography.

The current version of SURFACE II (version 3.0) consists of more than 100 routines written in Fortran-77, plus a small number of machine-dependent routines that perform functions not available in Fortran or that interface with file-handling and input-output devices (*10*). The program is controlled by a command language in which a four-character command word (such as GRID or STER) is followed by several parameters. If parameter values are not specified, the program assigns default values. The commands need not be given in any specific order, as SURFACE II will rearrange them into the proper sequence upon execution. However, if a nonstandard sequence of operations is required, this can be achieved by inserting PERF (Perform) commands into the sequence. This causes all preceding commands to be executed prior to reading any following commands. SURFACE II executes as a batch program, as do almost all other mainframe contouring programs. However, some of the programs have interactive (usually menu-driven) front ends that allow the user to build the file of commands more easily.

Gridding Algorithms in SURFACE II

The major differences between gridding-type contouring programs are in the way in which values of the surface are estimated at grid nodes. All algorithms, however, incorporate the commonsense notion that the value at a location on a surface should be closely related to values at nearby points and less closely related to values at distant points. Beyond some limiting distance two points should be independent of one another, and changes in the value at one location should have no effect on the value at the other. This gives rise to the concept of a neighborhood or zone of influence beyond which control points do not affect the estimation of the surface at a central location.

Many different schemes have been programmed to estimate values at the grid nodes. In the moving average procedure, an estimate is simply a distance-weighted average of values at control points within the neighborhood around a grid node. In effect, control points are projected horizontally to the grid location and the projections averaged. A more elaborate variant of this method is called the piecewise linear least-squares procedure: a plane that represents the "general slope" of the local surface is fitted by least squares to values at the control points within the neighborhood. The equation of the plane is evaluated for the grid location, which projects the values at control points parallel to the general slope. These projections are weighted and averaged. Sometimes a quadratic surface rather than a plane is used to represent the general surface.

A procedure called linear projection is used widely in commercial contouring programs. The local slope of the surface at each individual control point is first estimated by fitting a least-squares plane to other nearby control points. The individual fitted planes are then projected to the grid location and averaged. Linear projection is computationally more expensive than other gridding procedures because two sets of neighborhoods must be determined, first around each of the control points and then around each of the grid nodes.

Perhaps the most elaborate of all methods is punctual kriging, a statistical estimation procedure in which a set of simultaneous linear equations must be solved. The coefficients in the equation set are taken from a semivariogram, which describes the increase in differences between pairs of points as the distance between them is increased. The semivariogram is related to the spatial autocorrelation function and must be determined for every map. It is in effect a tailor-made weighting function for a specific data set. Since the locations of control points will be different within every neighborhood, the set of simultaneous equations must be solved for every grid node that is estimated. A typical contour map may have over a thousand grid nodes; if 16 control points are selected from within each neighborhood, a thousand sets of 17 simultaneous equations must be solved to make a contour map by punctual kriging.

Because SURFACE II was designed in part to investigate the relative advantages of different methods of contouring, it incorporates most of these alternative gridding procedures. Each procedure will produce contour maps with specific characteristics. For example, the moving average algorithm will create a surface in which the highest and lowest features coincide with control points, since an average cannot lie outside the range of values used in its calculation. The linear projection algorithm will extrapolate beyond the data, with control points lying on the slopes, but may produce unwarranted highs and lows on the edges of a map.

Figure 2A is a structure contour map of the southern part of the Alberta Basin, in central Canada. Contours are drawn on top of the Basal Fish Scales, a Cretaceous black shale that is a conspicuous marker horizon in records from oil wells. The unit has a pronounced regional dip toward the southwest. The map made by a moving average algorithm correctly reflects the regional dip near groups of control wells, but assumes a flattened, average attitude between wells. In contrast, a map made by using the linear projection algorithm correctly shows the pronounced regional dip, even in areas of low control (Fig. 2B). Both maps were made with ten nearest neighbors and a maximum neighborhood size of 25 miles.

Other Parameters in Gridding Algorithms

Within the neighborhood around a grid node, the influence exerted by a control point on the estimate at the node declines with increasing distance. Thus contouring programs weight nearby points more heavily in the calculation of an estimate. The form of the weighting function usually is an inverse power of distance. SURFACE II provides a selection of weighting functions ranging from d^{-1} to d^{-6}.

A contouring program may find a large number of control points within a neighborhood, most of them relatively far from the grid node being estimated. These will have almost no influence on the estimate and can be ignored. SURFACE II allows the user to specify the maximum number of points that will be used in the estimating equation, from as few as 4 up to as many as 48 points.

Almost any gridding algorithm will work reasonably well if the control points are uniformly distributed. However, if observations are closely spaced along widely separated lines, as they are in reflection seismic surveys or shipborne bathymetric surveys, an estimation procedure that simply uses the n closest points may produce unacceptable contour maps. Grid nodes near a line of observations will be estimated from the points along that line, whereas grid nodes halfway between two parallel lines of observations will be estimated by using points from both lines. The resulting map may resemble a contoured waffle, a consequence of the repeated abrupt changes in the selection of data from one line or the other.

This problem can be ameliorated by imposing radial constraints on the selection procedure so that points are selected in quadrants or octants around the grid node being estimated. Some nearby points (which have more information about the value of the surface) may be bypassed in favor of distant points that satisfy the radial constraint. This unfortunately produces a smoothed surface in which local details are averaged out, but it does reduce the worst effects caused by certain arrangements of control points.

Options in SURFACE II permit selection of a fixed number of control points from within a neighborhood, chosen with a quadrant or octant constraint or simply by distance. Alternatively, all points within a fixed

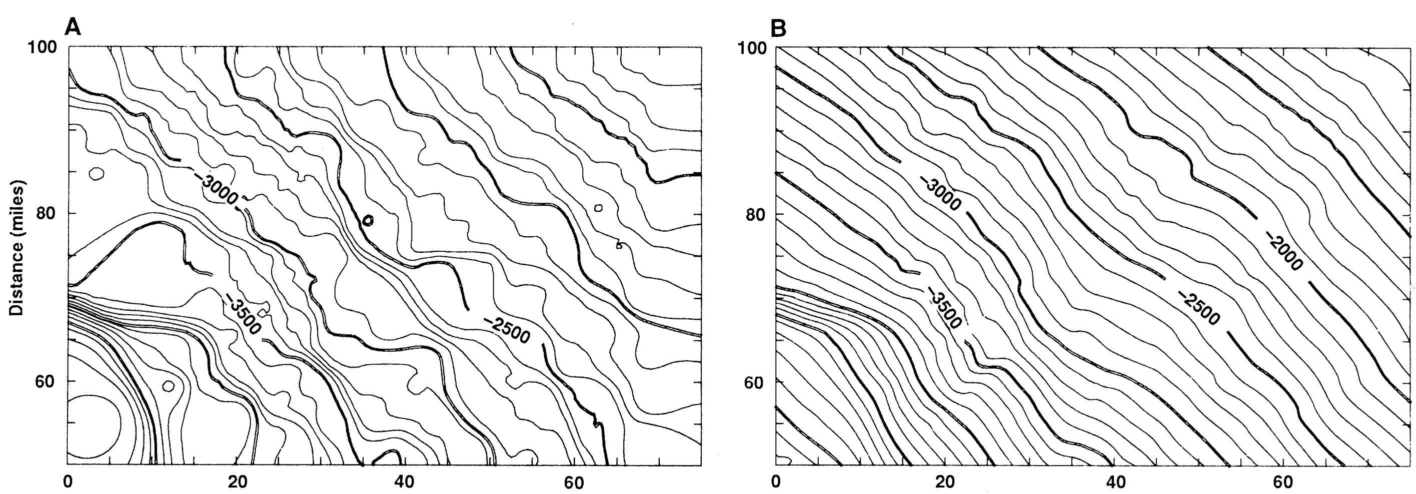

Fig. 2. Structural contour maps of Basal Fish Scales (Cretaceous) in southern Alberta, Canada, based on measurements in 360 exploratory holes. The contours are in feet below sea level. The map in (**A**) was constructed with a moving average algorithm, whereas the map in (**B**) was constructed with a linear projection algorithm.

Table 1. Data from SURFACE II error analyses of map grids used to construct Fig. 2. Errors were found by estimating values at control points from surrounding grid nodes and then by subtracting these from the original values. Error values are in feet except for the percent relative error.

Measure of error	Moving average		Linear projection	
	Error $x - \hat{x}$	Absolute error $\lvert x - \hat{x}\rvert$	Error $x - \hat{x}$	Absolute error $\lvert x - \hat{x}\rvert$
Maximum negative error	−42.602		−35.262	
Maximum positive error	30.259		25.990	
Mean error	−0.728	6.295	−0.293	3.672
Root-mean-square error	8.847	8.847	6.128	6.128
Standard deviation	8.817	6.216	6.121	4.905
Variance	77.745	38.643	37.462	24.064
Percent relative error	0.013	0.007	0.006	0.004
Skewness	−0.223	1.736	−0.785	2.812
Kurtosis	5.193	7.246	10.280	13.217
Sum-of-squares error	27,988.129	13,911.516	13,486.227	8,662.875

radius of the grid node can be used.

Statistical Analyses

SURFACE II contains many unique features for analyzing the original data and its contoured representation. It will perform statistical analyses of the spatial pattern of the control points and even produce maps showing various spatial relations among the data points. An error-analysis option will back-calculate from the grid to control points and determine if there are mismatches. A variety of statistics are calculated for these errors; this is one of the primary tools for evaluating the performance of a particular combination of gridding parameters. Table 1, for example, contains statistics taken from the error analyses of the maps shown in Fig. 2. Other options produce descriptive statistics and histograms of the original variables.

Surface Manipulation

The gridded surface can be manipulated in a variety of ways by SURFACE II. A filter option modifies the map grid by an arbitrarily weighted moving average to either smooth the surface or to emphasize specific characteristics. Similar filtering procedures are included in some commercial programs to improve the conformance of the gridded representation to the original data. Other options fit a polynomial regression, called a trend surface, in which the observations are the dependent variable and powers of the control point coordinates are the independent variables. SURFACE II will produce a contour map of either the trend surface itself or the residuals from the fitted surface. Another option will map an approximation of the local first derivative of a surface by calculating the slopes of the surface patches within the grid.

Applications

Contouring programs were developed to meet the needs of geologists, geophysicists, and geographers, but their use extends to most branches of science. In addition to applications in chemistry, such as that shown in Fig. 1, SURFACE II has been used by a Japanese shipbuilder to map stress fields in steel plates and by an archeologist to show the variation in the density of debitage at an archeological site. The program has also been used to display the distances that minority children had to travel to schools in an urban school district and to make contour maps of body temperatures measured by a medical scanner. The usefulness of a contouring program is limited only by the ingenuity of the user.

REFERENCES AND NOTES

1. Graphic by J. F. Remark, in R. J. Sampson, *Surface II Graphics System* (Kansas Geological Survey, Lawrence, 1978). For other examples see R. O. Gandour, G. M. Maggiora, R. L. Schowen, *J. Am. Chem. Soc.* **96**, 6967 (1974); J. F. Remark, *J. Chem. Educ.* **52**, 61 (1975).
2. Reviews of early developments are given by J. W. Harbaugh and D. F. Merriam, *Computer Applications in Stratigraphic Analysis* (Wiley, New York, 1968) and R. F. Walters, *Bull. Am. Assoc. Petr. Geol.* **53**, 2324 (1969).
3. C. E. Moore and B. Grethen, Eds., *Exploration Software and Data Sources for Micros* (Houston Geological Society, Houston, 1986).
4. M. J. McCullagh, *Geogr. Anal.* **13**, 51 (1981); D. F. Watson and G. M. Philip, *Math. Geol.* **16**, 779 (1984).
5. G. Birkhoff and L. Mansfield, *J. Math. Anal. Appl.* **47**, 531 (1974).
6. P. J. Green and R. Sibson, *Comput. J.* **21**, 168 (1978).
7. M. J. McCullagh and C. G. Ross, *Cartogr. J.* **17**, 93 (1980).
8. The CEED II (Comparison, Evaluation, Exhibition, Demonstration) meeting was held 19 to 22 August 1986, in Denver, CO. Results of the competition were reported in *Geobyte* **1** (no. 5), 25 (1986).
9. R. J. Sampson, *SURFACE II Graphics System* (Kansas Geological Survey, Lawrence, 1978); J. C. Davis and M. J. McCullagh, *Display and Analysis of Spatial Data* (Wiley, London, 1975).
10. Information about the availability of the SURFACE II software may be obtained by writing to Robert Sampson at the Kansas Geological Survey. SURFACE II is currently being revised; version 3.0 should be released by the end of the year. An interim version, with on-screen help but no user documentation, is now in use at the Kansas Geological Survey. As with earlier versions of the program, the source code will be available to academic institutions at a substantial discount. In addition, currently licensed users will be able to obtain the upgrade at nominal cost.