ASPECTS OF UNBIASED AND BIASED
CONTOURING OF GEOLOGICAL DATA
BY HUMAN AND MACHINE OPERATORS

by

E. C. Dahlberg

Amoco Canada Petroleum Company Ltd.
444 - 7th Avenue S.W.,
Calgary, Alberta

ABSTRACT

Representative contouring of geologic map data seeks to estimate the most likely configuration of an incompletely known surface from a sample of spatially distributed data points. Each contour line is, in effect, a statistic since it represents an estimate of a "true" value at an explicitly located geographic point.

By contrast, interpretive contouring presents a visualization of a conceptually meaningful geologic form such as a barrier bar, stream channel, dune or beach which the data could conceivably reflect. Each contour line is now a biased estimate based on speculation and additional information such as regional fabric and lithology, constrained only by the control points.

Thirty versions of a twelve point model map contoured by experienced professional geologists demonstrate the astonishing degree of variation attributable to just human differences.

Quantitative analyses of these maps suggest that (1) Areas of major ambiguity are mathematically predictable as a function of control point configuration. (2) Machine contoured maps closely approximate the unbiased "consensus" map. (3) The degree of interpretive license displayed by maps of the same data contoured according to delta front, salt dome, reef and channel, stream erosion and other models can be compared to and isolated from random operator variation.

Statistical methods for exposing subjective and objective aspects of contour maps serve geologists who are asked to accept hypotheses based on map evidence (in the scientific sense) and managers who bet cash on them (in the economic sense).

INTRODUCTION

A very large portion of the information communicated between not only geoscientists but members of other professions as well is done so in a graphic mode, i.e. as some kind of map. Many of these maps are based on numerical data which for the sake of clarity is contoured. A quantitatively contoured map is thus basically a picture of a surface which attempts to illustrate how some parameter of interest (generally a measurement or a count) varies over a planar area.
In the earth sciences a map is most simply a graph of a function \( z = f(X,Y) \) where \( z \) represents the magnitude of the mapped variable which varies over a portion of the Earth's surface as a function of \( X, Y \) which are latitude and longitude locators.

Specifically, then, a geochemical map pictures variation (inferred) of some chemical elemental content of the soil, rocks, etc. in an area. A structural contour map depicts spatial variation in the vertical position of the top of a formation of interest. Paleontological maps show percentage or abundance variation in fossil species, facies maps illustrate areal variability in mixtures of rock types and even phase diagrams are "maps" showing variability in the compositions of chemical compounds as a function of pH or pOH, temperature and pressure etc. Examples appear in Figures 1, 2 and 3.

THE NATURE OF QUANTITATIVE MAPS

These and all other "maps" have one basic point in common to wit that it is generally impossible and always impractical to obtain a measured value for the mapped quantity \( z \) for every possible combination of \( X \) and \( Y \). Thus it is necessary to estimate the behaviour of the mapped parameter over the major portion of the map area, so that some degree of uncertainty is incorporated in the final product.

The actual map itself is a planar picture of a generally non-parametric*, extremely complex surface such as the "carbonate isolith surface," the copper in parts per million surface," or the "surface of the sea floor," whose geometry is represented by patterns of contour lines. The problems therefore relate to aspects of accurately estimating and picturing such a surface from a finite collection of data points.

The major concern for the map user, be he a fellow scientist, manager, military man etc., is the evaluation of how faithfully a contoured map depicts the actual surface (be it hypothetical or real) from which the sample of points has been obtained. This paper examines such aspects as bias, subjectivity, accuracy, precision and objectivity and their roles in the process of quantitative contouring as well as evaluation of the resulting maps.

CONCEPTS

The following concepts have served as guidelines in our approach to the subject. Though proposed prior to the study, results of the experimental work have tended to reinforce our acceptance of their axiomatic utility.

1. Two mutually exclusive modes of quantitative data contouring are proposed, representative and interpretive.

*Polynomial approximations are sometimes utilized for representing the surface parametrically.
2. Representative contouring involves an attempt to depict the most likely configuration of an incompletely known surface (a) from a sample of spatially distributed data points.

3. Contouring in the representative mode is effectively unbiased, the contours are geometrically "best" estimates of the surface parameter values.

4. The errors are largely random thus the quality of the maps can be assessed by statistical criteria.

5. The other mode of contour mapping is interpretive; the purpose is to fit a conceptualized version of the surface into the data control.

6. Contouring in this mode is biased since estimates of surface values are largely governed by prior experience related to the phenomena concerned or additional data independent information.

In order to investigate critical aspects of human and machine contouring procedures a twelve point model map was contoured by thirty experienced professional geologists ... all people who make their livings with maps. The following points are evident.

1. An astonishing degree of variation is attributable to just human differences with no conceptual bias. Paired examples show perpendicular trends, areal "highs" on some are "lows" on others etc.

2. Variation in point estimates among the maps reflects ambiguity i.e., uncertainty and is a direct function of control point density, distance and magnitude. Uncertainty is thus theoretically mappable for purposes of quality control.

3. The "average" map represents the consensus of thirty estimates of the configuration of the "true" surface, so it constitutes the statistically best criteria for evaluating individual map patterns.

4. The machine contoured version of the same map is a close approximation to the consensus map suggesting that such (often controversial) maps may serve as effective criteria for documenting "interpretive license" in maps contoured by human operators.

5. The separation of subjective from objective aspects of contour maps is apparently possible by use of simple statistically appropriate methods. The details of these will be elaborated upon by this author in future communications.
FIGURE 1: Two types of contour maps. Left is an Isopach map showing how thickness of a sandstone varies regionally. Note extreme degree of crenulation in contours where data control (circles) is dense. This demonstrates relationship between degree of resolution of features and control point density. Right is a topographic map representing a portion of the earth's surface.
FIGURE 4 - This model map was used to evaluate comparative degrees of variation in human and machine contouring both from map to map and between various portions of the same map. The data are estimates of the elevation (below sea level) of the top of some geologic unit. The objective of representative contouring is to depict what the most likely form of the surface is, from the collection of data points provided. Some typical versions follow.
FIGURE 5 - Two contrasting though equally legitimate (with respect to the data control) estimates of the "actual" unknown but sampled surface. High degree of parallelism in R with two elongate NW-SE trending "highs." V shows a low on the west edge where R is high. Different regional trends.
FIGURE 8 - Two extremely contrasting versions particularly with respect to the locations of highs and lows. Note a 600 foot "high" to the left of center in N is a 900 foot depression in AA. Good demonstration of the relative magnitude of variation in estimates which can occur without violating the control.
FIGURE 9 - Z is supposedly a more valid version of what the actual surface really was. Its configuration was verified by more detailed drilling than has been presented on the model maps. Degree of map accuracy would be assessed by comparison with this map while degree of map precision would be assessed by comparison with the average or consensus map.
FIGURE 10 - Interpretive maps by Mr. Doug Lavoie of AMOCO Canada. Contouring is purposely biased by a geologic concept which "fits" within the framework of the data. Much of the map is highly speculative.
FIGURE 12 - Additional geological interpretations of the map data. In reality the interpretation would be tempered by independent considerations of regional geologic environment, lithology of units involved, etc.
FIGURE 13 - The computer generated map was contoured by a machine in an unbiased manner according to pre-specified, basically geometric rules. The sampling grid shows the points at which all maps were analyzed in order to arrive at the average and ambiguity maps.
MAP OF REGIONAL AMBIGUITY
(STANDARD DEVIATION AT EACH SAMPLE PT.)

FIGURE 14 - The consensus map is the average of all thirty versions of the model map. It is thus the best statistical estimate of what the "actual" surface really looks like. The ambiguity map shows areas of the map where variation in estimates of the actual surface was the greatest. A relationship to the density and proximity of control points is apparent.