A FRAMEWORK AND METHODS FOR CHARACTERIZING UNCERTAINTY IN GEOLOGIC MAPS

Donald A. Keefer
Illinois State Geologic Survey, Champaign, IL 61820, USA, keefer@isgs.uiuc.edu

1. INTRODUCTION

Geologic maps are products of complex analyses and interpretations, and as such, contain a certain amount of uncertainty. Unfortunately, it is difficult to estimate the uncertainty within a geologic map. Much of this difficulty is due to the lack of a clear framework for describing the sources of uncertainty within the map and poor awareness, within the mapping community, of methods capable of estimating these sources of uncertainty. Mann (1993) provides a discussion of uncertainty in geology and a framework for identifying the sources of uncertainty. His discussion is not limited to maps, but focuses more on data. Bardossy and Fodor (2001) use Mann's (1993) framework and extend his ideas by proposing four groups of mathematical methods that can be applied to characterize uncertainty in geology. Their discussion provides some innovative suggestions of methods for estimating uncertainty. As with Mann's (1993) treatment, they focus most of their discussion on data issues and do not provide any recommendations for conducting a comprehensive evaluation of uncertainty within geologic maps. I will present a framework for characterizing the sources of uncertainty in geologic maps which builds significantly on the insights provided by Mann (1993). I will extend Bardossy and Fodor's (2001) discussion by presenting a larger suite of practical methods for characterizing the uncertainty in geologic maps due to these different sources. Finally, I will briefly discuss concerns for conducting a more comprehensive, integrated evaluation of uncertainty in geologic maps and for interpreting and applying the results from such a comprehensive uncertainty analysis.

1.1 Objectives

In this presentation, uncertainty is defined as the difference between a predicted value and the real value. Uncertainty, therefore, is synonymous with error. A framework for understanding uncertainty within geologic maps is proposed which includes four major sources of uncertainty:

- data accuracy and precision,
- the amount and spatial distribution of data,
- the complexity of the geologic system being mapped, and
- geologic interpretations.

Identification and estimation of these separate sources of uncertainty is important, because each affects different aspects of mapping, and their collective estimation will lead to a more informed characterization and estimation of the total uncertainty of any geologic map.

2. UNCERTAINTY

The proposed component of uncertainty in geologic maps that is due to data errors is specifically related to the accuracy or precision of observations, measurements or calculations. Data errors affect what information and what interpretations can be reliably identified from the data. This proposed category combines Mann's (1993) categories of Observation Errors and Measurement Errors, and some of the situations covered by his Propagation of Errors category. Bardossy and Fodor (2001) note that the broad categories of probabilistic, possibilistic and hybrid methods all contain more-specific techniques that are appropriate for estimating errors within geologic data sets. Within these broad categories, the interval method and the use of fuzzy numbers are the most compatible with the kinds of error present in geologic data and with the limited knowledge that is generally available for quantifying most errors in geologic data. It is necessary to meaningfully estimate the errors within the data that are used for mapping to accurately estimate the impacts of these data errors on the overall accuracy of the maps and on any subsequent problem solving decisions.

2.1 Uncertainty due to Data

The component of uncertainty that is due to the amount and spatial distribution of data affects what features can be reliably identified within mapping areas and the accuracy of map-unit boundaries. This type of uncertainty will be affected by any data errors. This proposed category combines the Sampling Errors and the Errors of the Mathematical Evaluation of Geological Data categories of Mann (1993). Geologic maps are generally based on data sets that have non-uniform, or clustered, distributions of data. This spatial clustering can prevent the consistent identification of map features across the map area. There are a few methods that are useful for estimating or characterizing this type of uncertainty. These methods include the area of influence analysis (Figure 1), a non-traditional application of cross validation and specific analysis of the results from data-conditioned stochastic simulation.

2.2 Uncertainty due to geological complexity

The component of uncertainty in a geologic map that is caused by the complexity of the geology within the mapping area affects both the detail that is resolvable from each data type and the scale and fraction of actual geologic features that are identifiable within the map. Some geologic deposits are more complex than others and are more difficult to accurately identify and understand with the available data. This proposed category of uncertainty is essentially Mann's (1993) category of Inherent Natural Variability, and is a consequence of the underlying true distribution of the property of

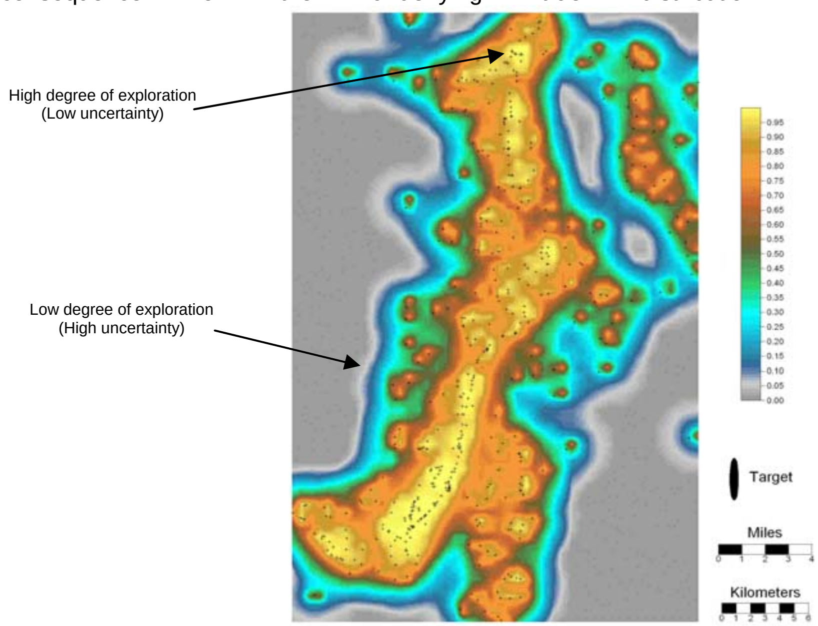

Figure 1. The probability that the target will be identified by the data points, if it is present. These probabilities are calculated using the area of influence method (Singer and Drew 1976). The data points are shown by the black points.

interest. This type of uncertainty is unaffected by data errors, amounts or distributions of data, or by our ability to identify and understand the actual distribution and characteristics of the geologic property. However, the ability to accurately estimate this type of uncertainty is affected by data errors, sparsely distributed or variably clustered data, and errors in geologic interpretation. Bardossy and Fodor (2001) recommend variability as the quantity that is used to evaluate geologic complexity. There are, however, many different ways to evaluate variability. Because the complexity of a geologic deposit, property or setting will change over a mapping area, methods are needed that can subdivide a mapping area into distinct regions and then the desired measure of complexity within each region can be calculated. If mapping is being done to address a specific application, it may be appropriate to select a measure of complexity that will be specifically relevant to the intended application. A few methods are available to estimate complexity based on the amount and structure of variability. These include: basic exploratory spatial data analysis measures (e.g., sample statistics, histogram analysis, transition probabilities, h-scatter plots, classified posting maps) which can be applied to 3-D regional subdivisions (e.g., Figure 2) or to small 3-D moving windows that can have a range of shapes or orientations; semivariogram analysis (Figure 3) which can be applied to larger 3-D regional subdivisions; and, a non-traditional application of cross validation which could also be applied to larger subdivisions.

2.3 Uncertainty due to Interpretation

The component of uncertainty that is caused by errors in interpretation is the most difficult to evaluate because it is composed of several factors which are interrelated in complex ways. This proposed category of uncertainty is a combination of several of Mann's (1993) categories, Conceptual and Model Uncertainty, Errors in the Mathematical Evaluation of Geological Data, and Propagation of Errors. Reliable estimation of the uncertainty in a map that is due to poor interpretations will be affected by all three of the other types of uncertainty that have been considered. Reliable estimation of this type of uncertainty will also require consideration of the types of interpretations that are made during a mapping effort, the order in which these interpretations are made and the way errors in earlier interpretations propagate to later interpretations. The types and interpretations that are made during a mapping project are often in the following sequence:

- defining the geologic framework of a mapping area,
- correlating observations, measurements and calculations from the data to the defined map units within the framework,
- interpolating properties between data locations, and
- finalizing interpolations for the end map products.

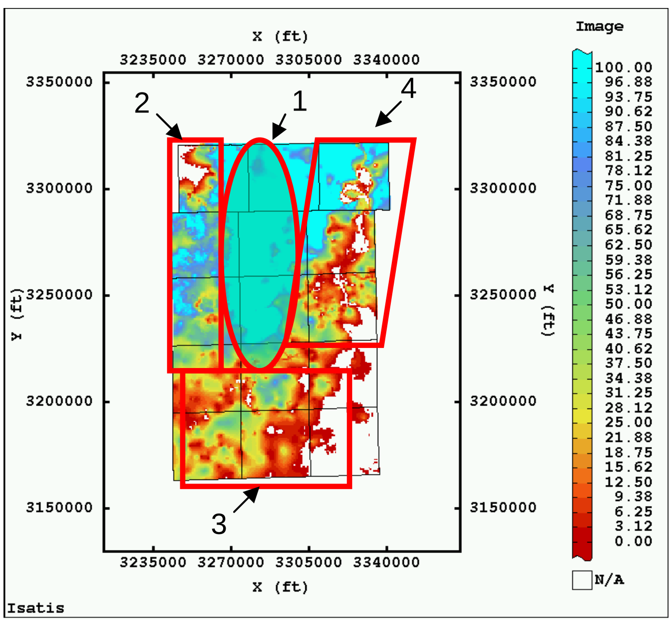

Figure 2. Regional subdivisions of the mapping area for evaluation of a specific mapping unit. The four numbered regions represent areas of distinct geologic complexity. These regions were defined using a range of properties, including available borehole data, geophysical profile data, land-surface geomorphic evidence, and conceptual models of suspected depositional environments.

While there can be iterations from step 4 back to step 2 or 3, this won't affect the dependency of the final interpretations on earlier ones. There are several tools for evaluating and estimating the uncertainty in a map due to errors in interpretation. They include, calculation of residuals between the data values and the mapped values; comparison of statistical characteristics between any combination of interpreted data, digital geologic map, conceptual models and modern or outcrop analogues; explicit description and delineation of the depositional, sedimentological and geomorphological conceptual models, including the expression of key properties like anisotropy, length scales and proportions; semivariogram analysis and comparison of calculations between interpreted data and geologic maps/conceptual models/modern analogues/outcrop analogues (Figure 3); analysis of results from data-conditioned stochastic simulation; and most importantly, an evaluation of the other 3 components of uncertainty and the potential implications they pose to the different interpretations.

3. SUMMARY

In approaching an uncertainty estimation and characterization effort for any geologic map, the proposed framework can help ensure that all the components of uncertainty are evaluated and their potential inter-dependencies are considered. Selection of specific methods for characterizing and estimating the sources of uncertainty within a geologic map should be based on an analysis of several criteria, which can include:

- mapping objectives,
- the size of map area, geologic complexity, depth of mapping and availability of data,
- intended map products,
- questions to be addressed by uncertainty estimations,
- geologic expertise of end users, and
- possible long-term uses of the maps,

This proposed framework and suite of methods provides many options to geologists. Even the independent evaluation of several different techniques can increase the relevance of the resultant uncertainty assessment for map users and decision makers. Estimation and characterization of the total uncertainty of a map can be realized based on the selection of appropriate methods and inclusion of reports that provide insight relevant to the mapping objectives, final products and expected application needs.

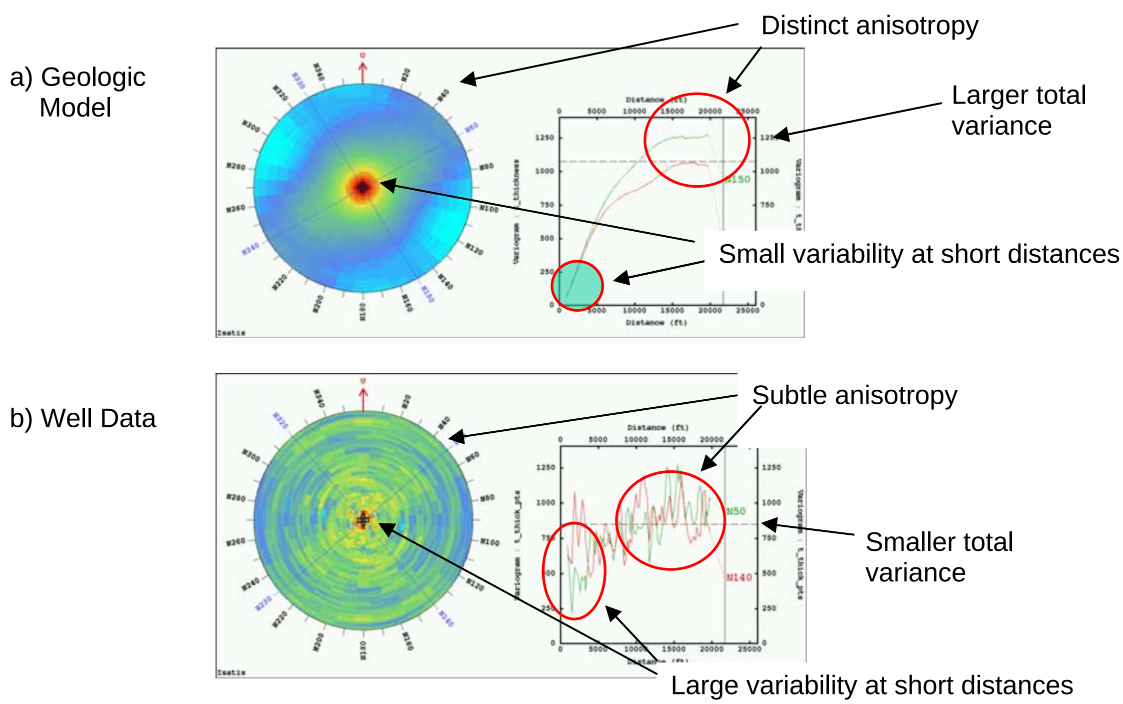

Figure 3. Semivariogram maps and directional semivariograms for one property of one map unit. Figure 3a includes a semivariogram map and directional semivariograms calculated from the draft geologic map. Figure 3b is the semivariogram map and directional semivariograms calculated using the map unit interpretations from the available borehole logs.

4. REFERENCES

Bardossy, G. and Fodor, J. 2001. Traditional and New Ways to Handle Uncertainty in Geology, Natural Resources Research, v. 10, no. 3, pp 179-187.

Mann, C.J. 1993. Uncertainty in Geology, In J.C. Davis and U.C. Herzfeld, (Eds), Computers in Geology – 25 Years of Progress, Oxford University Press, Oxford, pp. 241-254.

Singer, D.A. and Drew, L.J. 1976. The Area of Influence of an Exploratory Hole, Economic Geology, v. 71, pp. 642-647.