Geological reasoning and the problem of uncertainty

George Bárdossy (h4750bar@helka.iif.hu) and Janos Fodor

Basic concepts

Geological reasoning is a basic tool of all geological activities, but more than in other sciences, it is accompanied by uncertainties. Our goal is to review these problems and to present new mathematical approaches to handle them.

A major difficulty is that the terms of this subject are defined in very different ways. This is valid for the definition of the uncertainty itself. We suggest to apply the following simple definition: *Uncertainty is a lack of certainty in describing an object, a feature, or a process.* The following main types of uncertainties can be distinguished in mathematical respect, according to Dubois and Prade (2000) and Zimmermann (2000):

1) *Imprecision* or *inaccuracy,* expressing the deviation of measurements from a true value. We call the numerical value of this difference *error.*
2) *Vagueness* or *ambiguity* is the uncertainty of non-measurable objects and properties.
3) *Incompleteness* is the uncertainty due to incomplete information, that is, when the available information is not sufficient to perform the required mathematical procedure.
4) *Conflicting evidence* is the uncertainty arising from contradicting evidences, present in the given system.
5) *Presumption* or *belief* , when only general experiences are available about the given system.

In this context objective information (measurements, observations and descriptions) should be always distinguished from subjective information, based only on presumptions and experiences.

In geological reasoning the uncertainties are even more complex and require a more detailed classification. In our opinion, the lack of such a classification is the reason for most misunderstandings in geological reasoning. We suggest to distinguish at least the following main sources of geological uncertainty:

1. *Variability (heterogeneity),* a natural property of all geological objects, features and processes. The degree of variability, expressed by different statistics („measures of dispersion") can be used as a measure of this type of uncertainty. In geological systems both structured and unstructured variability may occur. *Structured variability* shows some regular spatial or temporal properties and can be described by mathematical methods, e.g. trend-surface-analysis. *Unstructured variability* occurs unexpectedly in a geological object and its spatial position and magnitude cannot be exactly predicted. The higher is the proportion of unstructured locations in a geological object, the larger is its overall uncertainty.
2. *Uncertainties due to imperfections of the geological investigations.*

2.1. *Lack of representative sampling.* It comprizes the inadequate volume of each sample, the choice of the sampling pattern, the sampling density, taking into account the „ranges of influence" and finally the inadequate size of the sample set.

2.2. *Inadequate choice of the laboratory measurements.* Often important features are not analysed to economize the expenses.

2.3. *The errors of the laboratory measurements,* including imperfect sample preparation, and homogenisation, calibration errors, imperfection of the instrument and the method of measurement, and incomplete skill and attention of the measuring personnel, finally confounding the absolute and relative errors.

2.4. *Uncertainties in the description of non-measurable properties* of the given geological object.

2.5. *Uncertainties due to subjective informations* (the expert's opinion)

2.6. *Uncertainties in the estimation and prediction of past and future geological events or processes.*

2.7. *Uncertainties of the geological modeling* (scale, parameter, and genetic models)

2.8. *Uncertainties of the mathematical modeling* (handling of extreme values- „outliers") etc.

2.9. *Uncertainties due to the incorrect application of the mathematical methods*

2.10. *Uncertainties in making final conclusions about the results of the given investigation..*

Natural variability, being a property of Nature, can be described and quantified by mathematical methods, but it cannot be diminished. On the other hand, the uncertainties of the geological investigations are consequences of human shortcomings and lack of adequate knowledge. They can be diminished to a certain extent, but they cannot be completely eliminated.

Mathematical methods for the handling of geological uncertainties

The type and the scale of the *input data* is of paramount importance for the mathematical evaluation of geological systems. The input data of the traditional deterministic and stochastic approaches are real numbers, that is, they do not express the uncertainty of their values. We call them „crisp data". It is easy to understand that only natural variability can be expressed by crisp input data, this being a limitation to the evaluation of the entire uncertainty of the given system.

The probability theory, being the base of traditional uncertainty assessments, comprises further limitations for the evaluation of uncertainties: The axioms of Kolmogorov, defining the theory of uncertainty, particularly that of additivity acknowledge only mutually exclusive events, that is, disjunct subsets. However, in geology disjunct subsets, such as objects and features are rare, transitions are much more frequent. Thus when applying the Kolmogorov axioms, we have to designate sharp boundaries where transitions occur. Obviously, this is a distortion of the natural reality, leading to biased results.

Even more inconvenient is the requirement of *repeated trials (experiments)* for most statistical calculations. Note, that the drilling and sampling of a bore-hole is an experiment in the statistical sense. In the case of a set of boreholes repeated experiments would mean to repeat the entire set of bore-holes after a shifting and rotating of the drilling locations. Obviously, such a procedure is unfeasable. In this and other similar cases the statistical calculations cannot be performed in correct mathematical sense. Furthermore, uncertain propositions and statements and subjective probabilities cannot be evaluated in terms of repeated experiments,thus they are a source of additional uncertainties.

Several geological features are not exactly defined and can be described only in a semi-quantitative or qualitative way. The probabilistic methods are not suitable for a quantitative mathematical evaluation of these data.

Geostatistics, based on the theory of regionalized variables (Matheron 1971) was an essential step ahead in the spatial evaluation of geological data sets, particularly the development of variography and kriging. However, geostatistics works also with crisp input data. For this reason it can evaluate only the uncertainty due to natural variability. Other theoretical limitations and shortcomings of geostatistics have been discussed by Diehl (1994) and Henley (2001).

For the reasons discussed above, we consider that the probabilistic methods are mathematically correct, but they cannot offer the optimal solution for several geological applications, particularly for the study of uncertainties. The *Bayes's theorem* offers a better approximation of these problems and it found a broad application in the last decade for the uncertainty analysis of engineering problems (Aven, Kveloy 2002). In our opinion, an even better solution can be achieved by the different *„uncertainty oriented methods"*developed by theoretical mathematicians in the last decades. The crisp input data of the traditional methods are replaced by special uncertain numbers.

Interval analysis (Moore 1979) replace the crisp numbers by „uncertainty intervals". It is assumed that the true value is somwhere within the interval. Interval analysis lacks gradations and is the simplest method to express uncertainty through arithmetic calculations. The method garantees that the true value will always remain within the interval, but this goal is achieved at cost of the precision. During the calculations the intervals become wider and wider and the final results become too conservative.

Possibility theory, a generalization of the interval analysis, provides a suitable model for the quantification of uncertainty by means of the possibility of an event (Zadeh 1978, Dubois,Prade 2000).The membership value of a number, varying between zero and one, expresses the possibility of the occurrence of that number. The related *fuzzy set theory* expresses uncertainty often by the use of fuzzy numbers. They represent estimates of uncertainty at different levels of possibility. Fuzzy numbers are by definition unimodal and they have to reach at least in one point the possibility level one, that is, the full possibility. In geology mainly trapezoidal and triangular fuzzy numbers are applied. They can be both symmetrical and asymmetrical. The smallest and the largest possible values of the given variable represent the lower and the upper bounds of the fuzzy number.All values of the variable must be within these boundaries. The values reaching the possibility level one are considered as the most possible estimates of the given variable. The fuzzy numbers are generalizations of the crisp numbers, as the latter ones can be regarded as a fuzzy number with a single point support.

All arithmetic calculations can be carried out with fuzzy numbers. One of their great advantages is that they do not require the knowledge of the correlations among the variables and the type of their probability distribution. For the sake of numerical comparisons and ranking, fuzzy numbers can be converted into crisp numbers. This calculation is called defuzzification.

We found that the *fuzzy set theory* (Zadeh 1965, Zimmermann 1996) is the most suitable, because its simplicity and flexibility, for geological applications. We performed with this method in the last years a number of successful test calculations in different fields of geology, such as quantitative mineralogical phase analysis by X-ray diffractometry (Bárdossy et al.2001) and by thermoanlysis (Földvári et al.2002), resource estimation of solid mineral deposits (Bárdossy et al. 2003), safety assessment of radioactive waste disposal (Bárdossy,Fodor 2001, Fodor,Bárdossy 2002) and paleontology (Bárdossy et al.2003). Further unpublished test calculations were performed on the evaluation of hydrogeologic measurements in boreholes (tansmissivity) and on some geomechanic parameters.

The methods of fuzzy logic have been applied by Cagnoli (1998) in volcanology.The development of *fuzzy geostatistics (fuzzy variography and fuzzy kriging)* by Bárdossy A. et al.(1990a and b) allows a much broader application of this method for the handling of uncertainties.

We agree with Zimmermann (2000) that no single uncertainty theory can claim to model all types of uncertainty, particularly in geology. For this reason we recommend to apply the traditional probabilistc and Bayesian methods together with the above discussed new, uncertainty oriented methodologies. In a recent paper Guyonnet et al. (2003) suggested a hybrid approach for combining Monte Carlo analysis with the use of fuzzy numbers. We are deeply convinced that even the general, theoretical problems of geological reasoning should be revisited by a joint application of the traditional and the new mathematical methods with a special emphasis to uncertainties.In this context we emphasize the evaluation of *linguistic descriptors,* omitted so far from the geomathematical evaluations.Note that the *membership functions* of the fuzzy set theory are capable to handle not only the uncertainties of geological reasoning, but also the various transitions of the geological systems.

A further important topic for geological reasoning is the field of *geological risks* in general, and of *natural hazards* in particular. The notion of risk has been defined in many different ways leading to much confusion. For us risk is the potential for the realization of unwanted and unexpected consequences of a decision or an action. The handling of risks by the traditional methods is mathematically correct , but by far not optimal. Here again we stress the necessity of the joint application of the traditional and of the new methods with special reference to the uncertainties of the statements of risk analysis. Biased or erroneous results of risk analysis may have fast developing consequences on economy and on large masses of local population.

As it was emphasized by Guyonnet et al. (2003), if a very large number of iterations is used in Monte Carlo random sampling scenarios with low probability will be realized, but with very low relative frequencies, due to the multiplication rule for independent events. As a consequence, these scenarios will be eliminated because they fall within the region of outliers. On the other hand, when uncertain parameters are represented by possibility distributions, these low-likelihood scenarios might not have been discarded, because the fuzzy calculation does not transmit through multiplication the uncertainty of the parameter values to the calculation results.

The quantitative assessment of the *reliability of risk statements* is of paramount importance and should accept particular attention. We intend to present our opinion and suggestions on this subject in the near future in a separate paper.

References

Aven,T., Kvaloy,J.T., 2002. Implementing the Bayesian paradigm in risk analysis. Reliability Engineering and System Safety.78.,195-201.

Bárdossy,A.,Bogárdi,i.,Kelly,W.E.,1990a. Kriging with imprecise (fuzzy) variograms I.Theory. Mathematical Geology.22.,63-79.

Bárdossy,A., Bogárdi,I.,Kelly,W.E.,1990b. Kriging with imprecise (fuzzy) variograms II.Application. Mathematical Geology.22.,81-94.

Bárdossy,Gy., Árkai,P., Fodor,J., 2001. Application of the fuzzy set theory for the quantitative phase analysis of rocks by X-ray diffractometry. Földtani Közlöny.Budapest. 131(3-4),331-341.

Bárdossy,Gy., Fodor,J., 2001. New possibilities for the evaluation of uncertainties in the safety assessment of radioactive waste disposal. Acta Geologica Hungarica.Budapest. 44 (4) 363-380.

Bárdossy,Gy., Kecskeméti,T., Fodor,J. 2003. L'emploi des méthodes de la théorie des ensembles flous a l'étude biométrique de l'espece Nummulites millecaput dans la partie occidentale de la Téthys. Comptes Rendus. Académie des Sciences. Paris. (in press)

Bárdossy,Gy., R.Szabó,I., Varga,G., 2003. A new method of resource estimation for bauxite and other solid mineral deposits. BHM. 148 (2)

Cagnoli,B.,1998. Fuzzy logic in volcanology. Episodes, 21, 94-96.

Diehl,P., 1994. Classifying geological uncertainty by geostatistical methods – many questions – few answers. Proc.Workshop on Reassessment of Coal and Mineral Deposits under Market Economy Conditions. Berlin. 164-175.

Dubois,D., Prade,H., 2000. Fundamentals of fuzzy sets. Kluwer Academic Publishers. Boston, London, Dordrecht, (647pp.)

Fodor,J., Bárdossy,Gy., 2002: Application of fuzzy methodology in the safety analysis of the Püspökszilágy radioactive waste repository, Hungary. Proc.3rd Int.Symposium „Computational Intelligence", Budapest 2002.Nov.14-15. 259-274.

Földvári,M., Bárdossy,Gy., Fodor,J., 2002. Application of fuzzy arithmetic to the quantitative phase analysis of rock samples using thermoanalytical methods, applied to the Boda Siltstone Formation,Hungary. Földtani Közlöny. Budapest, 132 (1), 1-15.

Guyonnet,D., Bourgine,B., Dubois,D.,Fargier,H., Come,B.,Chiles,J-P. 2003. Hybrid approach for addressing uncertainty in risk assessments. Journal Environ.Eng.Asce. 129 (1) 68-78.

Henley,S., 2001. Geostatistics – cracks in the foundations ? Earth Science Computer Applications, 16 (7) 1-3.

Matheron,G., 1971. The theory of regionalized variables and its applications. Cac. Centre Mrophologie Mathématique. Fontainebleau, (211pp.).

Moore,R.E. 1979: Methods and applications of interval analysis. SIAM Studies on Applied Mathematics. Vol.2. Philadelphia.

Zadeh,L., 1965. Fuzzy sets. Information and Control. 8. 338-353.

Zadeh, L. 1978. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems. 1. 3-28.

Zimmermann, H., J. 1996: Fuzzy set theory and its applications. Third ed. Boston

Zimmermann, H.,J., 2000. An application-oriented view of modeling uncertainty. European Journal of Operational Research. 122, 190-198.

Addresses

George Bárdossy, 1055.Budapest V.Kossuth-tér 18.,Hungary

János Fodor , Szent István University, Faculty of Veterinary Science, Department of Biomathematics, 1078.Budapest, István-u.2., Hungary, e-mail: jfodor@univet.hu