Rudarsko-geološko-naftni zbornik	Vol. 13	str. 15-24	Zagreb, 2001.

UDC 519.25

Izvorni znanstveni članak

UNCERTAINTIES AND RISKS IN GEOLOGICAL ACTIVITIES AND NEW WAYS OF THEIR HANDLING

Gyorgy BÁRDOSSY[1] and Janos FODOR[2]

[1] *Hungarian Academy of Sciences, Budapest - 1055. Budapest V.Kossuth-tér 18. Hungary*
E-mail: bar @mtrt.hu

[2] *Szent István University, Faculty of Veterinary Science, Department of Biomathematics, Budapest - 1078, Budapest, István u.2. Hungary*
E-mail:jfodor@univet.hu

Key words: Uncertainty, Errors, Geostatistics, Fuzzy sets, Risk analysis

The authors review the main types of uncertainties occuring in connection with geological investigations. The general concepts of handling these uncertainties are outlined. The particular features of scalar, spatial and spatial-temporal evaluations are presented. Limitations of the traditional mathematical approaches - deterministic and probabilistic - applied to geological investigations are discussed. In the second part of the paper new mathematical methods are presented that are more suitable to handle uncertainties than the traditional ones. The fuzzy set theory seems to be one of the most efficient for geological purposes, among these methods. The problems of uncertainty in risk analysis are shortly discussed. Finally, results of test calculations performed by fuzzy arithmetic are presented.

Ključne riječi: Nesigurnost, Greške, Geostatistika, Fazi-postupci*, Analiza (rudarskog) rizika

Autori prikazuju glavne oblike nesigurnosti geoloških istraživanja. U glavnim su crtama izloženi postupci za rješavanje tih nesigurnosti. Posebno su prikazani rezultati skalarne, prostorne i prostorno-vremenske obrade. Ocijenjena su ograničenja tradicionalnih pristupa interpretaciji rezultata geoloških istraživanja - determinističnog i probabilističnog. U drugom dijelu članka prikazane su nove matematičke metode, prikladnije od tradicionalnih za obradu interpretacijskih nesigurnosti. Čini se da je među tim metodama teorija fazi-postupka najefikasnija. Kritično su, ukratko, prikazani i problemi nesigurnosti u analizi rudarskog rizika. Konačno, prikazani su i rezultati nekih test-kalkulacija izvedenih fazi-aritmetikom.

* Od engl. "fuzzy", prekriven finim prahom (fuzz), dakle prašan, neproziran, nejasan, mutan, nesiguran.

Introduction

It is well known that all geological activities contain some types of uncertainty. They are mentioned in geological articles and reports, but surprizingly, little has been done so far to determine them and to take them into account (M a n n 1993). Even their definition is lacking from geological glossaries and therefore the terms of uncertainty and error are often confounded. In our opinion, *uncertainty* is the recognition that the results of measurements and observations may deviate more or less from natural reality. On the other hand, *error* is the difference between a true value and an estimate of that value. In this context *accuracy* expresses the closeness of the estimate to the true value. The term *bias* expresses the consistent under- or overestimation of the true value.

In pure scientific geological research uncertainties may result in false, spurious models, hypotheses and theories. The general scientific evolution sooner or later identifies them and leads to their rectification. In applied geology the consequences of uncertainties and errors may be quick and serious, sometimes even catastrophic: A mining investment based on false geologic assumptions will lead to an economic failure; a false calculation of landslide hazard may have catastrophic consequences for a community; a false safety assessment of a radioactive waste repository may endanger all kinds of life in a region. Consequently, it is highly reasonable to make efforts for a better understanding of uncertainties in geology, for their quantification, and if possible, for their reduction. The aim of this article is to review these problems and to offer new ways for their handling.

Types of uncertainties in geology

Two main sources of uncertainty can be distinguished in geology:

1. The inherent variability of Nature
2. The uncertainties of the geologic investigations

Natural variability is an inherent feature of all geological objects and processes. There are no completely homogeneous geological objects, even the minerals are not homogeneous, as their real crystal structure differs from the theoretical one. The degree of variability of a geological object may be highly different depending on the geological processes. There is a further important aspect: When studying the variability of geological objects, structured and unstructured features can be distinguished. The former ones are often called *trends*. Trends can be mathematically described by the well known methods of trend-surface-analysis. Spatial, temporal and combined four dimensional trend-surfaces can be distinguished. Examples of spatial trends are gradual compositional transitions of an igneous rock into another, or cyclic structures of some sedimentary sequences. On the other hand, on local scale unstructured features may occur unexpectedly and their spatial position and magnitude cannot be exactly predicted. There are no perfectly structured or unstructured objects and processes in geology, rather a mixture of the two types. The higher the degree of variability and the more it is unstructured, the larger is the inherent uncertainty related to the studied geological object.

The main uncertainties related to the *geological investigations* are due to incomplete knowledge and limited possibilities of the investigations. The main types of these uncertainties are as follows:

Sampling errors. The features of a geological object cannot be investigated at every possible point. Temporal and financial constraints allow only a limited number of bore holes, pits or trenches for sampling purposes. "Representative sampling", depicting in unbiased manner the features of a geological object, can be achieved very rarely.

Errors of field observations. Most geological features are observed in outcrops. Even these observations are often incomplete, due to unfavourable climatic conditions, dense vegetation cover or haste. Observations may be biased by lack of experience or by preferred personal interest and curiosity.

Errors of laboratory measurements. They consist of random and systematic components. The sources of measurement errors are: the imperfection of the instrument and of sample preparation (sample size, grain size, homogenisation, preferred orientation of some minerals etc), calibration errors, imperfection of the method of measurement, incomplete skill and attention of the measuring person. These problems have been amply discussed in mineralogy, geochemistry and engineering geology.

Errors due to temporal limitations. Geological processes can be directly studied at present time only and for a very short time interval, as compared with the length of geological periods. The more we go back in time the more uncertain are our conclusions about the nature and extent of the given geological process. The well known problem of *actualism* (uniformitarianism) adds a further component of uncertainty to the study of the geological past. The same is valid for the predictions of future processes e.g. in the case of safety assessments of toxic and radioactive waste disposal.

Conceptual and model uncertainties. When identifying and classifying geological objects, features and processes, existing geological concepts are applied necessarily. Are they always adequate to the given problem? Complete misidentifications may also occur. Natural analogues broadly applied to geological modeling are generally imperfect, as they cannot take into account undetected local features. Simplifications and generalisations made at geological modeling are further sources of uncertainty. *Scale models* (profiles, maps etc.) and *feature models* can be checked in a certain degree by the so called cross-validation methods, but for checking of *genetic models* only logical reasoning and geological experience ("experts opinion") can be applied.

When summarizing the sources of uncertainty in geology, it should be stressed that natural variability is a property of Nature, existing independently of us. On the other hand, the listed uncertainties of the geological investigations are due to imperfect human activity. In our opinion, conceptual and model uncertainties are the main sources of errors in geological investigations. A further adverse component is the subjectivity of scientific judgements. Erroneous theories and declarations of leading scientific authorities have been followed by the geological community at many cases for decades without criticism. It is enough to remind the rejection of the plate-tectonics theory for several decades.

General concepts of handling geological uncertainties

Despite the large number of uncertainties in geology there is a possibility to handle them more effifiently than before by applying the following successive steps:

a) Systematic consideration of uncertainties at each step of the geological investigation. This means a *new attitude* acknowledgeing that all geological activities comprise more or less uncertainty. At this first stage sources of uncertainty should be identified and their extent should be described by some sentences.

b) The uncertainties of the input data should be quantified, as well as possible. There is a number of new mathematical methods - to be reviewed later - allowing this quantification. We are aware of the existence of some unquantifiable features in geology, called *non-statistical uncertainties*. Let us take an example: Karts bauxite deposits are characterized by various forms of deposition, such as stratiform, blanket-type, lenticular, strip-like, graben fillings, canyon fillings, sinkhole fillings, pockets and nests. They represent a set of non-statistical features, since no probability distributions and other statistical parameters can be calculated on them. For these features only adequate descriptions of their uncertainty can be made. However, according to our experiences, the majority of geological object, features and processes can be quantified by the methods to be outlined later.

c) Having collected and adequate number of data and having quantified their uncertainties, a *mathematical evaluation* of this data set should be carried out. It should be stressed that no large data set can be quantitatively evaluated without the application of mathematical methods. On the other hand, any mathematical evaluation without the comprehension of the geological background is only an empty formalism.

The data taken into consideration can be divided into three groups regarding the degree of their uncertainty:

1) *Quantitative data*. They are results of measurements, their relative error being less than about 25%.

2) *Semi-quantitative data*. They are also results of measurements, but their relative error is more than about 25 %.

3) *Qualitative data*. They are the results of observations, expressed either by linguistic descriptions or by some verbal expressions of degrees, such as low, medium, high or very small, small, intermediate, big, very big etc.

In the past statistical evaluation of geological data was generally limited to quantitative, much more seldom to semi-quantitative data. The new mathematical methods, to be discussed later, enable us to evaluate qualitative data as well. This is of high importance, as according to our experience, qualitative data are very frequent in geological investigations.

All the above listed types of data must be evaluated in one of the following three frameworks:

A *Scalar evaluations*. These evaluations do not comprise the spatial and temporal position of the samples, only the sample values are of interest, e.g. statistical evaluation of the chemical and mineralogical composition of a geological object, identification of a set of fossils.

B. *Evaluation of spatially determined samples*. Each case (datum) has X,Y,Z spatial coordinates, e.g. exploration of mineral deposits, determination of the spatial position of tectonic structures.

C. *Spatially and temporally determined samples*. X,Y,Z coordinates and results of temporal measurements (time series) are related to each case, e.g. hydrogeological flow models, paleoclimatic reconstructions, safety assessments of radioactive waste repositories.

Each type of evaluation needs different mathematical approaches and different methods to determine the related uncertainties. Furthermore, the *scale of the study* must be also taken into account, such as micro-, macro-, regional- and global scale of study, as the geological processes may be quite different according the listed scales.

Finally, it must be emphasized, that the *incorrect application of mathematical methods* is a further source of error in geological investigations. The most common source of error is the insufficient number of sample elements collected. According to leading statisticians and theoretical results, the minimum number of a reliable statistical evaluation is about 30 samples. A further source of error is the neglect of mathematical rules, e.g. several statistical calculations require a normal distribution of the variable. Nevertheless, often variables with skewed distribution are evaluated by these methods. Evidently the results are biased.

Propagation of errors is also an important potential source of error. The errors related to the variables propagate differently depending on the interdependencies of the variables. Moreover, the propagation of errors through mathematical calculations is also different. With some methods it is absolutely necessary to avoid the repetition of the parameters in the calculations. For example, equation $a(b+c) = ab + ac$ holds only for real numbers. If a contains some error than the left side of the equation should be used instead of the right side where a appears twice. Neglect of error propagation may significantly increase the overall uncertainty. Even the computer programs ("codes") written for different geological models may contain erroneous assumptions, or programming errors.

Review of the traditional methods of mathematical evaluation in geology

When applying mathematical methods to solve geological problems, two types of approaches can be followed. The first is the *best guess or best estimate* approach, when so called point estimates are made, indicating what the author considers as the best result. Variability is expressed by some statistical parameters, such as variance and standard deviation, but no special attention is paid to quantify the uncertainties. The second is the *uncertainty oriented approach*. In this case, from the beginning of the investigation, at each step, the uncertainty is a matter of special attention. Even the initial *input data* should reflect their uncertainty.

The *deterministic methods* are the most common among the best guess approaches. They apply fixed, single parameter values, generally one of the measures of central tendency. Based on these parameter values, the geological object or process is described by differential equations. These methods are straightforward, but they do not take into account the possible error and the error propagation. However, in the absence of perfect knowledge of the given geological problem one cannot be sure about the exact values of the input parameters. All our calculations, although seamingly precise, harbor some degree of uncertainty. Consequently, deterministic methods can furnish unbiased results only if all the variables influencing the end results are known, their proportion is established and their interdependencies are known perfectly. In geological investigations, where perfect representative sampling is almost impossible and the influence of the different variables is only approximately detected, these requirements cannot be perfectly fulfilled. We consider therefore the deterministic methods as least suitable for geological applications, particularly for the study of uncertainties.

The *worst case analysis* (M o r g a n and H e n r i o n 1990) is an approach that notices the presence of uncertainty without modeling it explicitly. It works with the upper or lower bounds of a statistical distribution, trying to be sure that no larger or smaller value of the parameter may occur in the given system. This method is often applied for the safety analyses of radioactive waste repositories and in engineering geology. However, experiences showed that many worst case analyses produced hyper-conservative results.

The well known *probabilistic (stochastic) methods* found abroad application in geology. They operate with real numbers, called also *crisp numbers* as input data. The probability distributions and the different statistics, e.g. averages, measures of dispersion and skewness, are suitable to describe the errors due to natural variability, but their geological application has some theoretical limitations:

a) The additivity axiom - a fundamental law of the probability theory - recognizes only mutually exclusive populations. As a consequence, the method works with well defined boundaries between the populations (geological objects) and no transitions are admitted. However in geology sharp boundaries are rare, gradual transitions with mixed features are more frequent. Often there are more transitional zones in an area than pure objects.

b) Several statistical procedures require *repeated sampling*. In geological investigations this is mostly unrealizable. Imagine for instance a set of boreholes drilled in a regular grid. For the calculation of confidence intervals repeated sampling is required. This would mean repeated drilling of the grid after shifting and rotating the original drilling places. Obviously, such a procedure is unfeasable.

c) Closed systems, that is systems characterized by fixed sums, such as chemical analyses of rocks expressed by percentages, furnish biased results when calculating correlation among the variables. The methods suggested to rectify these errors are complicated and it is difficult to interpret their results (A i t c h i s o n , 1997).

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

For the above listed reasons we consider that the deterministic and probabilistic methods are mathematically correct for geological applications, but they are far to offer optimal solutions, particularly for the study of uncertainties.

Review of the main uncertainty oriented methods

A common feature of the uncertainty oriented methods is that they dissolve the above listed limitations of the traditional methods. Their most important feature is that they are able to describe mathematically the uncertainty of the input data by different types of *uncertain numbers*. Furthermore, they assure a correct propagation of the errors throughout all the calculations.

Interval analysis (M o o r e 1979) replaces the crisp numbers by uncertainty intervals (Fig. 1). The topic has become even more important with the advent of computers: the motivation is "the quest for rigor in numerical computation on machines". It is assumed that the true value is somewhere 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 precision. During the calculations the intervals become wider and wider and the final results become too conservative.

Possibility theory, a generalisation of interval analysis, provides a suitable model for the quantification of uncertainty by means of the possibility of an event (Z a d e h 1978, D u b o i s and P r a d e , 1988). The theory acknowledges that not all types of uncertainty can be handled by probability distributions. Instead, it uses membership functions to represent non quantified uncertainty. The membership value of a number, varying between zero and one, expresses the plausibility of the occurrence of that number. The theory has been applied successfully in biology, health

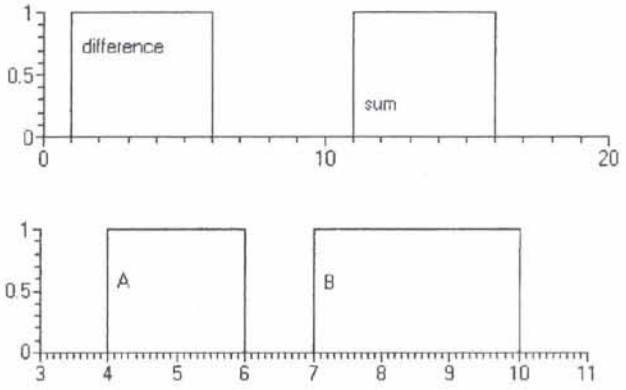

Fig. 1. Two intervals and their sum (A+B) and difference (A-B)
Sl. 1. Dva intervala te njihov zbroj (A+B) i razlika (A-B)

and medicine (Ferson and Ginzburg 1996, Ferson et al. 1999) and in different braches of industry and economy (Bárdossy A. and Duckstein 1995, Fodor and Roubens 1994).

The related *fuzzy set theory* expresses uncertainty very 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 (Takács and Várkonyi-Kóczy 1999a,b). For the sake of numerical comparisons and ranking, fuzzy numbers can be reconverted into crisp numbers. This calculation is called *defuzzification*. But the main advantage of the fuzzy method is that prior geological experience can be incorporated into the construction of fuzzy numbers. This goal can be achieved by joint constructing of the fuzzy numbers by geologists and the mathematician. The method allows the appropriate evaluation of semi-quantitative and qualitative input data as well. The frequent transitions of the geological populations, as mentioned before, can be also represented by fuzzy numbers (Fig. 2.). Cagnoli (1998) showed the application of the fuzzy set theory in the study of volcanic rocks. In the last years it found a broad application in the geographical information systems as well (Altman 1994, Macmillan 1995, Unwin 1995).

The way of constructing fuzzy numbers raises the problem of their *robustness*. Imagine that several well trained and experienced experts are asked to construct fuzzy numbers, based on the same crisp data. It is certain that the resulting fuzzy numbers will not be exactly identical. However, the differences are expected to be rather small. Luckily all the mathematical operations one has to carry out with fuzzy numbers are stable, that is, small changes in the input data yealds only small changes in the results. As a consequence, the final results are not sensitive to small differences in the initial fuzzy numbers.

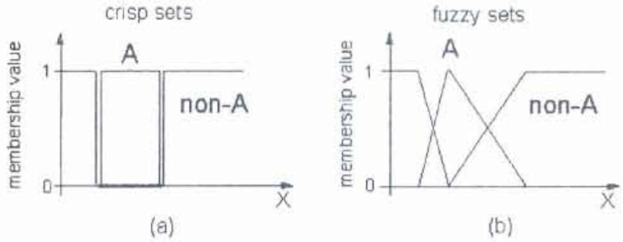

Fig. 2. (a) Crisp set A and its complement non-A. Their intersection is empty, and their union is the set of all elements of the universe
(b) Fuzzy set A and its complement non-A. They overlap
Sl. 2. (a) Jasan niz A i njegov komplement ne-A. Njihovo je sjecište prazno, a njihov spoj niz svih elemenata svemira
(b) Fazi-niz A i njegov komplement ne-A. Oni se preklapaju

The *probability bounds theory* (Ferson et al. 1999, Smith 1996, Teseem 1992) is a combination of probability theory with interval analysis. It expresses uncertainty by two cummulative probability distributions. The area between the two curves represents the extent of uncertainty of the given variable (Fig. 3.). Probability bounds are considered as a generalization of crisp numbers, intervals and probability distributions. The great advantage of this method is that it can apply different probability distributions, e.g. normal. lognormal, exponential etc. and correlations for the variables to be studied. But the method works also without making any prior assumptions. The probability bounds get narrower with more empirical information about the given geological object. Its disadvantages are the more complicated calculations to be carried out. Nevertheless it seems for us to be a highly efficient approach in the case of safety assessments, when prior information is abundant.

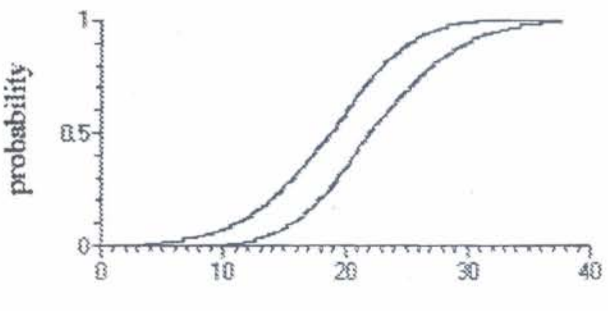

Fig. 3. Probability bounds
Sl. 3. Veze po vjerovatnosti

Neural networks (Aleksander and Morton 1990, Kosko 1992) build models directly based on measurements. They consist of adaptable nodes, which through a process of learning from task examples, store experimental knowledge and make it available for use. Neural networks are suitable to evaluate complex geological systems and processes that are too complicated to be understood by traditional modeling. A trained neural network can be thought as an "expert" and used to provide predictions for new objects or situations. The method, however, requires a large number of data about the object or process to be studied. Neural networks are particularly suitable for the investigation of potential mineral resources. In this case there is a great uncertainty in distinguishing potential deposit and non-deposit areas for later drilling. Singer and Kouda (1999) showed on test examples from the United States that this problem can be solved by neural networks. Recently *neuro-fuzzy systems* (Fullér 2000) were developed, enlarging the method by the aspects of data-uncertainty.

The method of *hybrid arithmetic* (Cooper et al. 1996, Ferson and Ginzburg 1996) combine probability distributions with intervals, fuzzy numbers and probability bounds. The method allows the use of all kinds of numbers, this being its greatest advantage. This is the newest among the methods of uncertainty analysis and there are very few publications on its application. Test calculations are foreseen by us for the near future.

All the above listed methods require for the statistical calculations at least about 30 cases as sample size. Below this number the results become more and more uncertain (Tukey 1977). This circumstance represents difficulties for many geological applications. The recently developed *bootstrap method* (Efron and Tibshirani 1993, Davison and Hinkley 1997) allows to diminish the sample size to about 10 cases by performing computerized random sampling and calculating the statistics of each replicate. Up to 1000 replicates can be quickly obtained by adequate computer programs. The method has been tested recently by Z. Sebestyén on a Late Permian geological formation situated in southern Hungary. The results proved the mathematical correctness and efficiency of the method. Another method developed for calculation with small sample sizes is the *jacknife* (Rock 1988). It is similar to the bootstrap method with the difference that it produces replicates by omitting one data in turn and then it averages the statistics of the trimmed replicates.

Calculations of spatial uncertainty

The calculations with the uncertainty oriented methods, outlined above, did not extend to spatial problems. However, in geology the spatial position of geological objects and features is of primary importance. For any spatial evaluation spatial coordinates (X,Y,Z) must be added to each input data and the spatial position must be part of all further calculations.

The theory of *reqionalized variables*, called also "geostatistics" (a term leading to many misunderstandings), developed by Matheron (1971) offered a method for spatial evaluations within the framework of the traditional probability theory. The calculation of **variograms** allowed to quantify spatial natural variability and to determine the "range of influence" in space for the given variable, that is the range of spatial autocorrelation. In our opinion, variograms are highly important tools to understand and to determine the spatial variability of any geological feature and to diminish their spatial uncertainty, e.g. interpolation between two boreholes is reasonable only within the range of influence, beyond it is no more than pure formalism. Spatial predictions can be carried out by *point* and *block-kriging* and even the error of the prediction can be determined by the *kriging standard deviation.* The results can be represented on isoline maps. Matheron's theory was a real break-through for spatial calculations in geology, mainly for reserve estimations and mining-geology, but it also has some limitations, such as the requirement of first and second order *stationarity* in the study area. A further limitation is, that semi-quantitative and qualitative input data cannot be evaluated by this method. Matheron's method has been broadly applied during the last decades in petroleum exploration and mining as well. Much has been published on the results that we do not intend to repeat here. Only the article of Schuenemeyer and Power (2000) should be mentioned, as they studied the uncertainty of coal resource assessment in North Dakota, USA, by separating large-scale and local effects of natural variability and applying semi-variograms of coal thickness.

In our opinion, Matheron's theory is the best solution for spatial averageing and for the determination of natural variability, but it does not take into consideration the uncertainty of the input data, discussed in the first part of this paper. The development of *fuzzy geostatistics* (Bárdossy A. et al. 1990a,b), particularly of fuzzy variograms and fuzzy kriging, was an essential step for evaluating spatial uncertainty in geology, including the uncertainty of the input data. In our opinion, spatial equivalents can be developed for all the non-spatial uncertainty oriented methods, discussed in the foregoing chapter. This is an urgent task for the coming years.

Calculations of spatial and temporal uncertainty

The study of temporal processes requires the introduction of another dimension, that of *time*. The traditional methods of time-trend analysis are well known and applied in geologic investigations (Davis 1986). New aspects of space-time relationships have been developed (Roddick and Hornsby 2001) and geostatistical methods have been expanded to space-time models (Kyriakidis, Journel 1999). Bárdossy A. and Duckstein (1995) applied fuzzy rule based models to time-series analysis of hydrogeological problems, e.g. water demand forecasting. Nevertheless, the problems of uncertainties and errors have been treated so far in other geological investigations only by traditional deterministic and probabilistic approaches. Here again further theoretical studies are needed to apply the uncertainty oriented methods to spatial and temporal evaluations. Spatial-temporal predictions have a particular importance in the safety assessment of radioactive waste disposal (Craig 1988).

Uncertainty of risk analysis in geology

Risk is a common term in science, economy and industry. According to the definition of the Society of Risk Analysis, *risk is the potential for realization of unwanted consequences of a decision or an action. Risk analysis* is defined by the same society as "the process of quantification of the probabilities and expected consequences of risks" (2001). Risk analysis has been applied to several problems in geology, such as mineral exploration, mining projects, landslides, floods, volcanic and earthquake hazards. The safety assessments of toxic and radioactive waste repositories represent particularly important applications of risk analysis. All these calculations have been carried out so far by traditional deterministic and probabilistic methods (Bonano and Cranwell 1988, Craig 1988, Hunter and Mann 1992). At our knowledge, no uncertainty oriented methods have been applied for risks of geological problems so far.

The basic requirement of risk analysis is to exclude the possibility of under-estimation of risk at the given conditions. With the traditional method measures of central tendency (mean, median etc.) are produced. However, experience showed that not these measures, but the tail of the distributions are of paramount importance, as they represent risks of low probability, but of severe consequences. *Dependency bounds analysis*, suggested by Ferson (1996), seems to assure sufficiently secure estimates of these tail-probabilities.

The methods of interval analysis and fuzzy arithmetic have been first applied to risk analysis by Ferson and

K u h n (1992) for ecological problems. Our aim is to apply these methods for the calculation of geological risks as well.

Test calculations

During the last year a number of test calculations have been carried out by us mainly with fuzzy arithmetic. We found this method as most suitable for the study of uncertainties in geological investigations. Examples of these applications are presented in the following:

Quantitative mineralogical phase analysis of rock samples is carried out mostly by X-ray diffractometry and by differential thermogravimetry. The methods are well known and have been described in many papers and textbooks. They complete each other and are applied generally jointly. We evaluated them separately for methodological purposes, outlined below. The *differential thermogravimetric method* is based on changes in weight and entalpy of the minerals, occurring during the heating of the rock sample. The measurements have been carried out by a MOM Derivatograph-PC. The heating occurred at 10°C/min speed and the TG, DTG and DTA curves were registered by the computer. The average error of the phase analysis was considered so far to be about ± 10 weight%, the limit of detection varies from 1 to 5 percent. It was clear that part of this error is due to the apparatus and the imperfection of detection and registration. But experience showed that the main source of error is the incomplete knowledge of the rock forming minerals. Let us stress that minerals showing no weight and entalpy changes during the heating, cannot be detected at all by this method.

The test calculations have been carried out on 27 rock samples taken from the Late Permian Boda Claystone Formation (BCF), southern Hungary, and on one bauxite sample, taken from the Szőc bauxite area in the Bakony Mountains. Fuzzy numbers have been constructed for every mineral of each sample. This process has been performed together by the expert of the derivatography and by the two authors of this paper. This is an essential point of all evaluations of this type, as the technical, mathematical and mineralogical-geological aspects of the evaluations must be discussed and decided together. The characteristic values of each fuzzy number, such as the possible minimum and maximum of the base (support) and the point or interval of the core, the fuzzy index and the defuzzificated numbers have been determined and compared with the crisp numbers, obtained by the traditional evaluation.

Figure 4. shows a comparison of the traditional and the fuzzy evaluation of an albitic claystone sample from the BCF. The fuzzy evaluation clearly shows how different are the errors of the different minerals. The results of the evaluation of the 27 BCF samples are presented on Table 1. Illite-muscovite could be detected with the largest error, followed by chlorite, calcite, dolomite and by montmorillonite. The lengths of the core (membership value one) in the average is much shorter than the support and this expresses the real error of the phase analysis. The sum of the thermally active minerals is only 45-65 weight % and the loss of weight is also very limited. Thus this is not a very suitable rock for thermal investigation. On the other hand, bauxites are generally very suitable being composed mainly of thermally active minerals and the loss of weight may reach 30 %. The sample from Szőc contains 81 % of thermally active minerals. As a consequence, the fuzzy evaluation revealed only relatively little errors of measurement, such as ± 2 % for boehmite, ± 1/2 % for gibbsite, ±1 % for goethite and ±6 % for kaolinite.

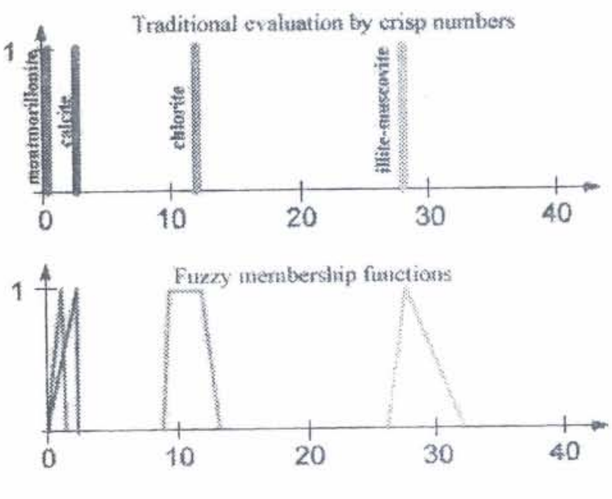

Fig. 4. Comparison of the evaluations by crips and fuzzy numbers. Thermogravimetry by derivatograph.
Borehole BAT – 10. 58.7 m, albitic claystone, BCF

Sl. 4. Usporedba vrijednosti putem jasnih i fazi-brojaka. Derivatografska termogravimetrija.
Bušotina BAT – 10, 58,7 m, albitski glinjak Boda-formacije

When summarizing the results, it can be stated that the averages of the traditional crisp numbers and those of the defuzzificated numbers are in very good agreement. Their standard deviation expresses the *natural variability* of the composition fo the BCF. On the other hand, the standard deviation of the corresponding fuzzy numbers expresses the *analytical error* of the phase analysis, carried out be derivatography. The latter is 2 to 4 times higher in the studied samples than the natural variability. On our opinion, this is a very important additional information, furnished by the fuzzy evaluation. The detailed discussion of the evaluation, including all the data obtained, is in press (F ö l d - v á r i , B á r d o s s y , F o d o r 2001).

The well known quantitative phase analysis by *X-ray diffractometry* was performed on 32 samples taken from the same formation (BCF) as before. The measurements

Table 1. Characteristic values of the fuzzy numbers calculated from 27 rock samples of the BCF
Tablica 1. Karakteristične vrijednosti fazi-brojki izračunatih na temelju 27 uzoraka uzetih u formaciji Boda-glinjaka u južnoj Mađarskoj

Mineral	Support %	$\bar{x}$ %	Core %	$\bar{x}$ %	Average of the fuzzy numbers %
1. illite-muscovite	1-29	16	0-24	6	23
2. chlorite	3-26	12	0-20	4	15
3. dolimite	1-11	5	0-4	1/2	9
4. calcite	1-14	6	0-6	1/2	6
5. montmorillonite	1-6	2	0-1	0	2

were carried out by Philips PW-1730 diffractometer, monochromatised CuK_{α} irradiation at 45 kV and 35 mA. Chemical analyses were obtained by X-ray fluorescence analysis on each sample. The accepted average error of the traditional phase analysis was ± 10 % and the limit of detection 1/2 to 5 weight %. First the crisp values of each sample have been determined by the traditional evaluation method (Bárdossy et al. 1980). In the second step the corresponding fuzzy numbers were constructed in the same way as with the derivatography - outlined above. A comparison of the traditional and fuzzy results of an albitic claystone sample is presented on Figure 5. The determination of each mineral was achieved with different errors of detection. The minerals can be ranked according to the amount of this error, as presented on Table 2. The sequence of the minerals is the same as that obtained by derivatography, completed by the thermally inactive minerals, such as albite, quartz and hematite. Here again the length of the coreinterval is much shorter than that of the support.

The averages of the mineral content, calculated separately from the crisp and the defuzzified numbers showed an almost perfect coincidence. They express - as in the case of derivatography - the natural variability of the BCF regarding the studied mineral. On the other hand, the standard deviations of the fuzzy numbers are 3 to 6 times higher than those of the crisp numbers, expressing the analytical error of the phase analysis.

The main benefit of the fuzzy approach is - with both types of measurement - that particular features of each mineral and each sample can be evaluated separately and the minerals can be ranked according to the error of their determination. According to our experiences, this approach can be applied to any other rock as well. A detailed description of the above outlined evaluation, including all the data obtained, is actually in press (Bárdossy, Árkai, Fodor 2001).

Reserve estimation of mineral deposits. The reliability of reserve estimates is a problem for more than 50 years.

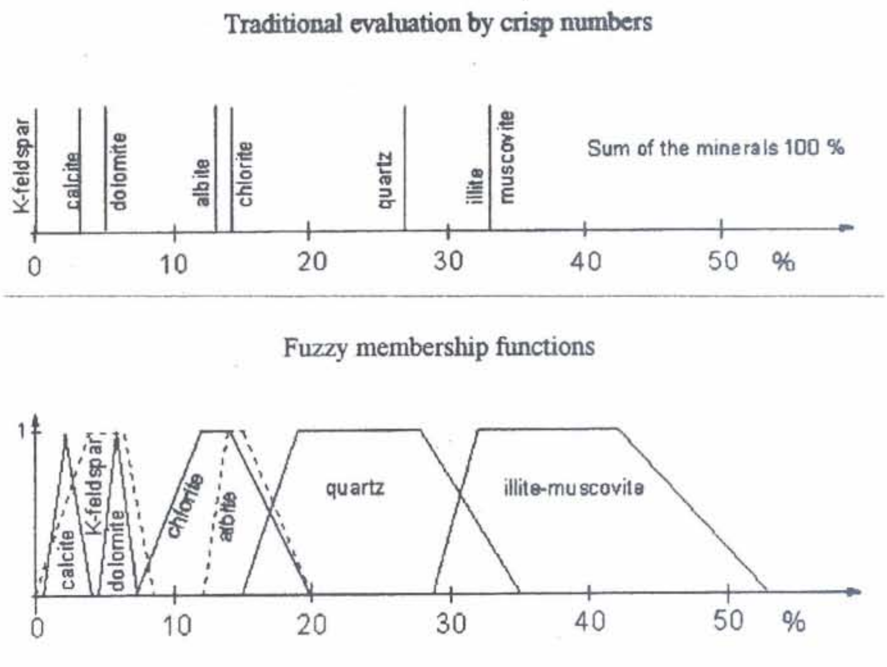

Fig. 5. Comparison of the avaluations by crisp and fuzzy numbers. X-ray diffractometry. Albitic claystone, BCF, Borehole BAT – 4, 1169.2 m depth

Sl. 5. Usporedba vrijednosti dobivenih putem jasnih odnosno fazi-brojki. Rendgenska difraktometrija. Albitski glinjak, Boda-formacija. Bušotina BAT-4, 1169,2 m

Table 2. Characteristic values of the fuzzy numbers calculated from 32 rock samples of the BCF

Tablica 2. Karakteristične vrijednosti fazi-brojki izračunatih na temelju 32 uzorka uzeta u formaciji Boda-glinjaka u južnoj Mađarskoj

Mineral	Support %	$\bar{x}$ %	Core %	$\bar{x}$ %	Average of the fuzzy numbers %
1. illite-muscovite	5-25	14	1-14	5	34
2. albite	2-25	12	0-16	3	31
3. quartz	2-16	7	0-12	4	9
4. chlorite	3-14	6	0-6	2	7
5. dolomite	1-13	4	0-10	2	8
6. calcite	2-10	4	0-6	2	6
7. hematite	2-5	3	0-3	1	5

Hundreds of articles have been published on this subject, but so far only deterministic and probabilistic approaches have been applied. The most important development in this field was the application of the theory of regionalized variables, called geostatistics. But even geostatistical reserve estimates apply traditional crisp input data. The method of fuzzy arithmetic has been applied by us first to a large bauxite deposit in Hungary (B á r d o s s y , et al. 2000). This work has been continued, in cooperation with the Bakony Bauxite Mining Co. by applying it to eight further bauxite deposits having different types of deposition. The main results are as follows:

In our opinion, the first step of any reserve estimation is a thorough evaluation of the *deposit model*. Even within the group of karst bauxite deposits, this can be very different, e.g. stratiform, lenticular, valley, graben, sinkhole, pocket, string, nests and all their combinations. In a second step the *ranges of influence* must be determined by variograms at least for the thickness of the commercial bauxite and its main chemical components. The fuzzy evaluation can be started only if the ranges of influence cover the entire deposit, not letting in between any undetectable spaces situated outside of these ranges. Fuzzy numbers are then constructed for the productive area, the average thickness and the average bulk-density of the deposit. Here again a close cooperation of the mathematician and the bauxite-geologist is of fundamental importance. For the *productive area* first the minimum and maximum values of the base (support) are determined. The maximum area is determined by a line connecting the improductive bore holes situated closest to the deposit. The minimum area is determined by the connecting line of the outermost productive boreholes. The most possible area - characterized by membership value one - has been constructed on the base of isopach maps and a close set of geological cross-sections. The above three points form a triangualr fuzzy number (Fig. 6/A).

The average commercial bauxite thickness is evaluated by the traditional statistical method, taking into account all the boreholes, where the bauxite thickness is more than the cut-off value: at present for this mining company ≥ 2,0 meters. According to our experience, the distribution of bauxite thickness is often very skewed in the direction of high thicknesses. This inevitably leads to the overestimation of the reserves. To eliminate this bias, instead of the mean, so called *M-estimators* are calculated, being robust, unbiased representations of the "central tendency". Tukey's M-estimator seems to be the most suitable for the karst bauxite deposits. The fuzzy number, expressing the average bauxite thickness, is constructed around this M-estimator, taking as core the "*standard error of the mean*". The minimum and maximum values of the base are calculated by the confidence interval at 95 % ($\alpha = 0.05$) level of confidence. The resulting fuzzy number is trapezoidal (Fig. 6/B). The values of the *bulk-density* are determined in the laboratory on relatively small samples, completed by some large (1/2 - 1 m^3) ones, measured in the mine. The construction of the fuzzy number, expressing the average bulk density, occurred in the same way as with the average bauxite thickness, resulting in a trapezoidal fuzzy number (Fig. 6/C).

The amount of *geological reserves* was calculated by multiplication of the above discussed three fuzzy numbers (Fig. 6/D). The resulting fuzzy number expresses the total error of the reserve estimation in tons. Let us stress that the minimum and maximum points of the base express the smallest and the largest possible amount of reserves, if all components of the reserve calculation are either favourable or unfavourable. Obviously, such a coincidence is completely improbable, but the fuzzy number expresses just the possible extreme values of the reserve estimation. On the other hand, the core involves the most acceptable reserve estimates, with equal plausibility. *There is no pre-*

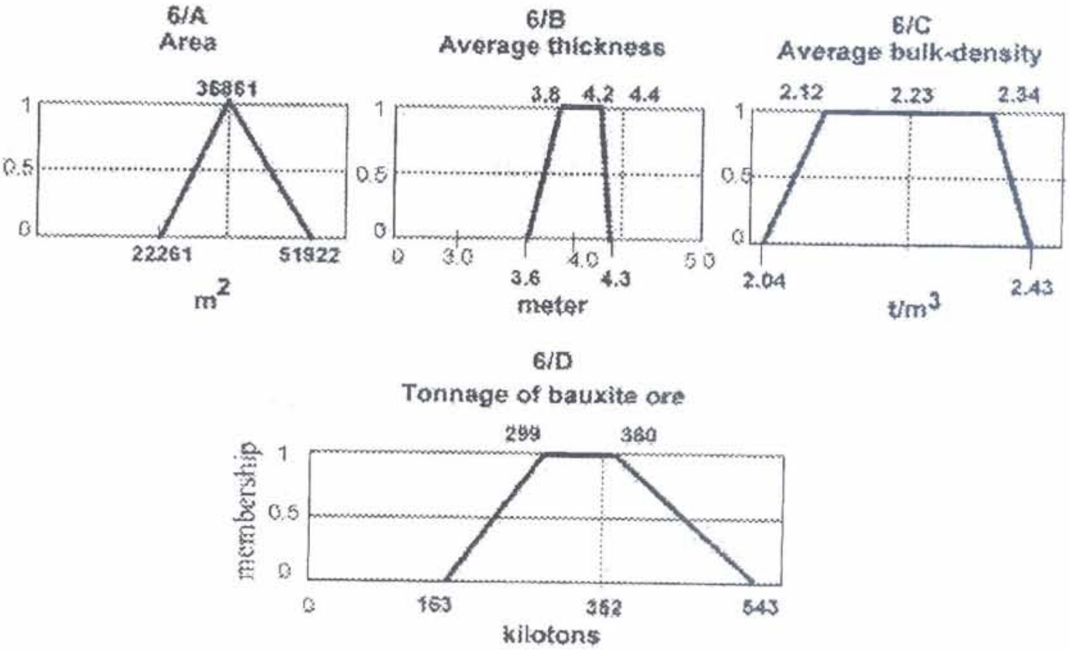

Fig. 6. Fuzzy numbers of the reserve estimation of the bauxite deposit Szárhegy, Hungary (Thin lines show results of the deterministic reserve estimation)

Sl. 6. Fazi-brojke u procjeni rezervi boksitnog ležišta Szárhegy u Mađarskoj (tanke linije pokazuju rezultate determinističke procjene rezervi) *ferred point within this interval, including the result of the traditional point-estimate!* The average quality of the ore is calculated in the same way, again taking into account the very frequent skewness of the distribution.

A great advantage of the fuzzy evaluation is its clearness and simplicity. One can easily recognize the weak points of the reserve estimate and the amount of related error. Based on the reserve and quality estimates, it is easy to calculate the risks of the given mining project. If the risks are unacceptably high for the investing company, complementary exploration can be started. As the procedure of fuzzy evaluation and risk analysis is easily performed by adequate computer programs, the entire calculation can be repeated after the drilling of each new borehole. Thus exploration can be stopped after reaching the required risk-level. This way exploration expenses can be minimized.

Further test calculations were carried out by us on hydrogeological (transmissivity) and rock-mechanical data, with positive results. The limited extent of this paper does not allow to discuss these applications. Maybe the most important new application of uncertainty oriented methods was performed by us for the *safety assessment of radioactive waste repositories*. A completely new methodology has been elaborated and suggested for application. Its detailed discussion is also in press (Bárdossy and Fodor 2001).

Conclusions

1. Uncertainty has long been considered in geology a removable adverse circumstance that should gradually disappear with the overall development of the Earth-Sciences. However, one must recognize that a part of this uncertainty is an inherent feature of Nature. Therefore, understanding and appropriate handling of uncertainties should be part of all future geological investigations.
2. The traditional mathematical methods - deterministic and probabilistic - applied so far in geological investigations are mathematically correct, but by far not optimal for the treatment of uncertainties.
3. The new mathematical methods, developed by theoretical mathematicians, are suitable to evaluate in a mathematically correct way semi-quantitative and qualitative ("linguistic") input data and to determine the uncertainties and errors connected with them.
4. In our opinion the traditional and the new uncertainty oriented methods complete each other in the geological investigations.
5. A thorough study of the geological objects and processes is indispensable for any mathematical evaluation in geology. Without that even the most sophisticated method becomes an empty formalism.

Received: 2001-05-24
Accepted: 2001-10-23

REFERENCES

Aitchison, J. (1997): The one hour course in compositional data analysis, or compositional data analysis is simple. In: Pawlowsky, V. (editor): Proc. Of the 3rd Annual Conf. of the Internat. Assoc. of Math. Geol., 3-35 pp.

Aleksander, I. & Morton, H. (1990): An introduction to neural computing. Chapman and Hall, 240 pp, London.

Altman, D. (1994): Fuzzy set theoretic approaches for handling imprecision in spatial analysis. *Int. Jour. of Geographical Information Systems, 8*, 271-289.

Bárdossy, A., Bogardi, I. and Kelly, W.E. (1990a): Kriging with imprecise (fuzzy) variograms I. Theory. *Math. Geol., 22*, 63-79.

Bárdossy, A., Bogardi, I. and Kelly, W.E. (1990b): Kriging with imprecise (fuzzy) variograms II. Application. *Math Geol., 22*, 81-94.

Bárdossy, A. & Duckstein, L. (1995), Fuzzy rule based modeling with applications geophysical, biological and engineering systems. CRC Press, 232 pp, New York.

Bárdossy, G. (1997): Some fields of geomathematics as seen by a geologist. In: Pawlowsky-Glahn, V. (editor): Proc. IAMG Conf. Vol. I, 36-56 pp., Barcelona.

Bárdossy, G., Árkai, P. and Fodor, J. (2001): Application of the fuzzy set theory to the quantitative phase analysis of rocks by X-ray diffractometry. Földtani Közlöny (in press).

Bárdossy, G. & Fodor, J. (2001): New possibilities for the evaluation of uncertainties in safety assessment of radioactive waste disposal. *Acta Geologica* (in press).

Bonano, E.J. & Cranwell, R. M. (1988): Treatment of uncertainties in the performance assessments of geologic high-level radioactive waste repositories. *Math. Geol., 20*, 543-565.

Cagnoli, B. (1998): Fuzzy logic in vulcanology. *Episodes, 21*, 94-96.

Cooper, J. A., Ferson, S. and Ginzburg, L.R. (1996): Hybrid processing of stochastic and subjective uncertainty data. *Risk Analysis, 16*, 785-791.

Craig, R. G. (1988): Evaluating the risk of climate change to nuclear waste disposal. *Math. Geol., 20*, 567-588.

Davis, J. C. (1986): Statistical and data analysis in geology. 2nd edition, Wiley, 684 pp, New York.

Davison, A. C. & Hinkley, D.V. (1997), Bootstrap methods and their application. Cambridge University Press, 582 pp, Cambridge.

Dubois, D. & Prade, H. (1988): Possibility theory: An approach to computerized processing of uncertainty. Plenum Press, 263 pp, New York.

Efron, B. & Tibshirani, R. J. (1993): An introduction to bootstrap. Chapman and Hall, 436 pp, New York.

Ferson, S. (1996): What Monte Carlo methods cannot do. *Human and Ecological Risk Assessment, 2*, 990-1007.

Ferson, S., Root, W. and Kuhn, R. (1999): RAMAS Risk Calc: Risk assessment with uncertain Numbers. Applied Biomathematics, 184 pp, Setauket, New York.

Ferson, S. & Ginzburg, L.R. (1996): Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety, 54*, 133-144.

Ferson, S. & Kuhn, R. (1992): Propagating uncertainty in ecological risk analysis using interval and fuzzy arithmetic. In: Zanetti P. (editor): Computer Techniques in Environmental Studies, Elsevier Applied Science, 387-401 pp., London.

Fodor, J. & Roubens, M. (1994): Fuzzy preference modeling and multicriteria decision support. Kluwer Academic Publishers, 272 pp, Dordrecht.

Földváry, M., Bárdossy, G. and Fodor, J. (2001): Application of fuzzy arithmetic to the quantitative phase analysis of rock samples by thermoanalytical methods, applied to the Boda Aleurolite Formation, Hungary. *Földtani Közlöny* (in press).

Fullér, R. (2000): Introduction to neuro-fuzzy systems. Physica Verlag, 289 pp, Heidelberg.

Hunter, R. L. & Mann, C. J. (1992): Determining probabilities of geologic events and Processes. Studies in Mathematical Geology No. 4, Oxford University Press, 306 pp, Oxford.

Kyriakidis, P.C. & Journel, A.G. (1999): Geostatistical space-time models: a review. *Math. Geology, 31*, 651-684.

Kosko, B. (1992): Neural networks and fuzzy systems. Prentice Hall, Englewood Cliffs, New Yersey.

Macmillan, W. (1995): Modelling: fuzziness revisited. *Progress in Human Geography, 19*, 404-413.

Mann, C.J. (1993): Uncertainty in geology. In: Davis, J. C. and Herzfeld, U.C. (editors): Computers in Geology - 25 Years of Progress, 241-254 pp.

Matheron, G. (1971): The theory of regionalized variables and its applications. Cah. Centre Morph. *Math. Fontainebleau, 5*, 211 pp.

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

Morgan, M.G. & Henrion, M. (1990): Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge.

Rock, N.M.S. (1988): Numerical geology. Springer, 427 pp, Heidelberg.

Roddick, J. & Hornsby, K. (editors) (2001): Proc. Of the International Workshop on Temporal, Spatial, and Spatio-temporal Data Mining. Lecture Notes in Artificial Intelligence 2007. Springer, Heidelberg.

Schuenemeyer, J. & Power, H.C. (2000): Uncertainty estimation for resouce assessment - an application to coal. *Math. Geol., 32*, 521-541.

Singer, D.A. & Kouda, R. (1999), A comparison of the weights-of-evidence method and probabilistic neural networks. *Natural Resources Research, 8*, 287-298.

Smith, J.E. (1996): Generalized Chebychev inequalities: theory and applications in decision analysis. *Operations Research, 42*, 807-825.

Takács, O. & Várkonyi-Kóczy, A.R. (1999a): Fuzzy handling of uncertainty in nonlinear systems. In: De Baets, B., Fodor, J. and Kóczy, L.T. (editors): Proc. of EUROFUSESIC '99 (Budapest, May 25/28, 1999), 22-27 pp.

Takács, O. & Várkonyi-Kóczy, A.R. (1999b): Information processing based on mixed-classical and fuzzy-data models, In: IEEE International Workshop on Intelligent Signal Processing (Budapest, September 4-7, 1999), 23-27 pp.

Tessem, B. (1992): Interval probability propagation. *Int. Jour. of Approximate Reasoning*, 7, , 95-120.

Tukey, J. W. (1977), Exploratory data analysis. Addisson-Wesley, 688 pp, Reading, Mass.

Unwin, D. (1995): Geographical information systems and the problem of error and uncertainty. *Progress in Human Geography, 19*, 549-558.

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