Earthquake Engineering, Tenth World Conference © 1992 Balkema, Rotterdam. ISBN 90 5410 060 5

Probabilistic treatment of uncertainties from incomplete knowledge in SHA

Yuxian Hu & Han-Yao Chen
Institute of Geophysics, State Seismological Bureau, Beijing, People's Republic of China

ABSTRACT: Sources and current methods of analysis of uncertainty from randomness,fuzziness and ignorance or incomplete knowledge in seismic hazard assessment problem are briefly discussed at beginning; understandings of the authors are then presented in the following order. All three types of uncertainty come from incomplete knowledge. Probablistic method can be applied to all of them,objective probability for random factors and subjective probability for the other two types of uncertain factors. Discrete subjective probability density functions for incomplete and fuzzy factors can be obtained from logic–tree and membership functions respectively. Fractile curves may be used to show the scattering of any uncertainty factor, but a unified probabilistic treatment may be applied to any combination of all three types of uncertainties.

1 INTRODUCTION

Since method of safety factor was replaced by probabilistic or reliability method in structural design, uncertainties of different natures were studied in many fields. Seismic hazard assessment (SHA) is a prediction or estimation, deterministic or probabilistic, of the future earthquake effect to be expected in a region or at a site. It is universally accepted that, compared with evaluation of static or wind loads, SHA involves much more uncertainties, random or others.

The probabiblistic method currently used was suggested by Professor Cornell in 1968(Cornell,1968), with the occurrence of earthquake treated as a random process, following Poisson distribution, for simplicity. The final result of a SHA will be a hazard curve, which gives various possible ground motion intensities corresponding to different probabilities of exceedance. This method has been widely applied in the world.

There are many other types of uncertainties involved in SHA in addition to randomness of the earthquake occurrence. The randomness of the relationship of earthquake magnitude and fault rupture length and the ground motion attenuation has been discussed in the 70's. In the 80's,uncertainties involved in SHA have been studied in detail during the SHA works done for the re–evaluation of the safety of nuclear power plants in the Eastern United States, because earthquakes are rare and their evaluation is more uncertain than that on the west coast. Uncertainties in modeling of both earthquake occurrence and attenuation, in potential sources zonation, and in estimation of seismicity parameters for each potential source have been discussed in detail. Early works emphasized the sensitivity analysis of the uncertainty factors (Zhu et.al,1980) and suggested that most part of the uncertainty came from attenuation(Kiureghian et.al,1977). Works in the last decade emphasized the differences in nature and, consequently, the methodology of analysis of these uncertainties. They (Bernreuter et.al, 1987; Coppersmith et.al, 1986; Hu, 1990; McGuire, 1987; Savy et.al, 1986) classified uncertainties in two types, random and ignorant or incomplete knowledge. Lind classified in his nice review (Lind,1985) the uncertainty factors in structural safety analysis in three types, namely, random, fuzzy and incomplete; he mentioned that Bayesian method can be applied to analyze jointly the random and incomplete uncertainties, but methods were not available to analyze jointly the fuzzy uncertainty with others.

Uncertainty analysis is of practical importance (Whitman,1989), but there is still a need of some unified methodology to analyze jointly uncertainties of different natures occurred together in one problem. The present paper suggests such a method which reflects clearly the current level of understanding on the knowledge, complete or incomplete, of the experts of that speciality.

2 SOURCES AND NATURES OF UNCERTAINTIES

McGuire (McGuire,1987) defined the random uncertainty as the inherent property of the event adopted the usual Bayesian method as follows

$$P[Y>y] = \int P[Y>y|\mu]\, f(\mu) d\mu \qquad (1)$$

where $f(\mu)$ is the probability density function of random factor μ for $0<\mu<\infty$; $P[Y>y|\mu]$ is the conditional probability of some ground motion parameter Y over some given value y when the random factor takes a value of μ. For ignorant factors,a logic–tree approach is widely accepted to express the belief of one group of experts in term of various possible choices of the uncertain factors and their combinations together with weights assigned to each of them. Fig.1 shows a sketch of such a logic tree.

Potential source zonation		Annual rate of earthquake occurrence		...	Branch Terminal
φ_{1k}	P_{1k}	φ_{2k}	P_{2k}	...	μ_1
		case1	0.3	...	P_1
case1	0.4	case2	0.5	...	P_2
		case3	0.2	...	...
		...		...	...
case2	0.5	...		...	...
case3	0.1			...	...
...				...	P_n

Fig.1 A sketch of logic tree in SHA

As shown in the sketch of Fig.1, uncertain factors φ_{jk} ($j=1,\cdots,m$; $k=1,\cdots k_j$), such as the boundary of the potential source φ_{1k}, annual rate of earthquake occurrence of the source φ_{2k} etc., are listed from left to right.Suppose there are three possible choices for one source zonation, cases 1 to 3 ($k_2=1,2$ or 3) with weights $P_{1k}=0.4,0.5$ and 0.1 respectively assigned by a group of experts. Under case 1 of the source zonation,three or more possible cases with weights $P_{2k}=0.3$, 0.5, $\cdots$, etc. and the final terminal of each branch is listed in the right column,with weight $P_i = \Pi P_{jk}$, which is the product of the weights P_{jk} of all uncertain factors φ_{jk} in that branch.The sum of the weights of all possible cases of one event must be one to satisfy the requirement of probability.Weight is just another name of subjective probability or degree of belief. The final uncertainty variable μ_i ($i=1,2.\cdots,m$) is only a dummy variable,which is but a sequence index but represents a combination of values taken by all uncertain factors φ_{jk} in a branch.

Each branch forms a Monte Carlo sampling of all uncertain factors considered, from which one curve of exceedance probability of SHA can be calculated by the routine mehod to obtain one sample of the curve $P[Y>y|\mu_i]$.

McGuire and others believe that the effect of uncertain factors given in the sketch of Fig.1 can be considered only by the fractile curves as shown in Fig.2. Their idea is that random uncertainty may be treated by Eq. (1), but incomplete uncertainty should be treated by fractiles of the Monte–Carlo SHA curves.

Savy et al. (1986) listed several methods to show the effect of the uncertainty factors, such as best estimate, fractile, arithmatic and geometric mean and pointed out the large differences between them, as shown in Fig.2.

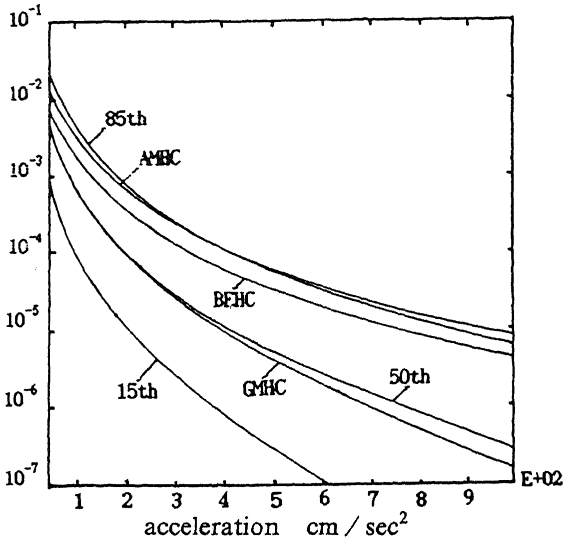

Fig.2 Effect of uncertainty from incomplete knowledge

Authors' suggestion is to use Eq.(1) to consider the effects of all three types of uncertainty, but the scattering curves of different choices can be used to show uncertainties still existing, which can be reduced by further studies. The only requirement is to show the adequacy of using a probability density function $f(\mu)$ to describe the uncertainty and to find it for fuzzy and ignorant factors.

4 PROBABILITY DENSITY FUNCTIONS OF UNCERTAINTY FACTORS

(1) Fuzzy Uncertainty

If the reversed question " what is the probability of a fuzzy factor being in a state x or taking a value of x under some given fuzzy information" is accepted, a subjective probability density function may be defined and obtained as follows. A fuzzy set A is defined by its membership function

$$\mu_A(x) = \mu_1 / x_1 + \mu_2 / x_2 + \cdots\cdots + \mu_n / x_n = \sum \mu_i / x_i \qquad (2)$$

where the plus and summation signs are not arithematic, but mean only assembly, and μ / x

is not fraction but a membership relation, i.e. member x has a right μ to belong to the fuzzy set A. $\mu=1$ means 100% right, $\mu=0$ means no right, and $\mu=0.2$ means 20% right. This is the universally accepted way in fuzzy set theory. Another widely accepted idea is fuzzy probability (Zou et.al, 1989)

$$P(A) = \sum P(x_i)\mu_A(x_i) \qquad (3)$$

where P(A) is the probability of the occurrence of fuzzy set A, $P(x_i)$ the probability of occurrence of the member x_i. The prior probability $P(x_i)$ may be taken as constant when there is no other information. From fuzzy Bayessian theorem

$$P(x_i|A) = \mu_A(x_i).P(x_i) / P(A) \qquad (4)$$

and the fuzzy probability P[A] may be written as

$$P(A) = P(x_i)\sum \mu_A(x_i) \qquad (5)$$

and finally, from Eq.(4), the probability density function of a fuzzy set A over its members x_i $(i=1,2,\cdots,n)$ is obtained

$$P(x_i|A) = \mu_A(x_i) / \sum \mu_A(x_i) \qquad (6)$$

which is definied subjectively by experts at discrete members $i=1,\cdots,n$. Accordingly, if the membership function of middle–aged person is

$$\mu_A(x) = 0/10 + 0.1/15 + 0.3/20 + 0.6/25 + 1/30 + 1/35 + 1/40 + 0.7/45 + 0.2/50 + 0.1/55 + 0/60$$

the subjective discrete probability density function of the middle–aged is then

$$P(x_1=10|A) = 0.0/5.0 = 0.0$$
$$P(x_2=15|A) = 0.1/5.0 = 0.02$$
$$\cdots$$
$$P(x_5=30|A) = 1.0/5.0 = 0.20$$
$$\cdots$$
$$P(x_N=60|A) = 0.0/5.0 = 0.0$$

which is shown in Fig.3. Here, the age axis takes discrete value at the center of the interval ($x_i - 2.5, x_i + 2.5$)only.

The final result for a dummy fuzzy factor μ with discrete interval $\triangle\mu$ is

$$f(\mu_i)\triangle\mu = P(x_i|A)$$
$$= \mu_A(x_i) / \sum \mu_A(x_i) \qquad (7)$$

and $\sum f(\mu_i)\triangle\mu = 1$. Because the membership function is subjectively assigned by expert, the descrite probability function thus obtained must be subjective.

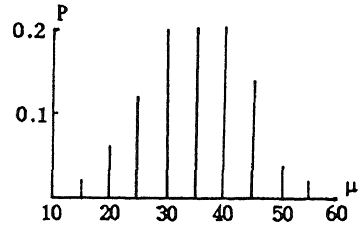

Fig.3 Discrete probability function of a fuzzy quantity

(2) Incomplete Uncertainty

As mentioned above, Monte–Carlo samples can be obtained from incomplete factors in a form of logic tree as shown in Fig.3 where the weights given at the branch terminals may be treated as a subjective discrete probability density function. The only difference is that the uncertainty variable μ is a combination of the values of all incomplete factors considered in the logic tree; it is a symbolic dummy variable; the function $f(\mu)$ is actually not a function in this case, but an assembly, and $f(\mu)$ and μ forms a pair of assembly.

(3) Monte–Carlo Sampling of Random Factors.

Monte–Carlo samplings can be generalized numerically from any given probability density function $f(\mu)$ of a random factor μ, such as ground motion attenuation. If the number of sampling is large enough, the samples of μ will be distributed as the given density function. A group of μ values thus generated can be used identically for the probability density function $f(\mu)$. For each sample of μ, a SHA curve can be calculated in the routine manner to obtain one SHA sample curve, quite similar to the logic–tree approach for the incomplete uncertainty. The only difference is that the probability density is objective and sometimes continuous for random factors, but always subjective and discrete for fuzzy factors, and usually a subjective assembly for incomplete factors when several such factors are considered simultaneously.

5 A UNIFIED APPROACH FOR ALL TYPES OF UNCERTAINTY

Once a probability density function, objective or subjective, is obtained for an uncertain factor, authors suggest that it may be considered in SHA through the Bayesian approach, i.e., Eq.1. For given discrete probability density function of a fuzzy factor μ, Eq.1 becomes

$$P(Y > y) = \Sigma P(Y > y|\mu)\, f(\mu)\triangle\mu \qquad (8)$$

for an incomplete factor of given branch terminal weights or subjective discrete probability P

which cannot be reduced by the completeness of the understanding of the physics of the event. Examples of the random uncertainty in SHA are the scattering of ground motion data in attenuation laws and the earthquake occurrence. McGuire and others considered uncertainties in modeling of attenuation and earthquake occurrence,potential source zonation,seismicity parameter evaluation as ignorant or incomplete uncertainty, which comes from the incompleteness of knowledge on the factor and can be reduced when the understanding of the factor is improving (Bernreuter et.al, 1987; Coppersmith et.al, 1986; McGuire, 1987; Savy et.al, 1986).

Lind(1985) added a third type of uncertainty to random and incomplete, i.e. fuzzy, which comes from human–being's simplification. According to authors' understanding, fuzzy ideas are subjective classifications in qualitative terms, such as young man, strong motion, and heavy damage. The indices, boundary, and definition of the classification are all not clear but fuzzy. Each person or expert has his own understanding of a fuzzy idea; only the common rough understanding of the fuzzy idea is important and the differences of individual's understanding are ignored. For example, old man may mean generally a man above 65 years old.

The normal problem of a fuzzy idea may have certain or uncertain answer. For example, a man of 40 years old has 100% right to belong to the group of middle– aged, but a man of 20 years old may have 20% right according to one and 30% right according to another to belong to that group. In reverse problem, which is neccessary in SHA, however, the answer is always uncertain. For example, a problem like "a young man is 20 years old" or "a strong motion is 0.2g" is not certain.

Authors believe that all three types of uncertainty come from incompleteness of knowledge on the subject.

Randomness of an event is known to come from the numerous number of factors, each of which has some but not decisive influence on the event and from both the unknown variation of these factors and the unknown way they affect the event. It is the authors' belief that, following the understanding of some of these factors, of the variation of them, and of the way they affect the event, the randomness of the event will be reduced. For example, after some further understanding of the distribution of the strength, rigidity, deformation and stress on a fault and in the crust, the occurrence of earthquake on that fault will be less random. Although these factors are too many and too difficult to understand now, they will be known gradually following the development of science and technology. Compared with 40 years ago, the development of the science of plate tectonics is certainly a big step forward to reduce the randomness of the understnding of the earthquake occurrence; it is not unreasonable to predict that after another 40 years the randomness of earthquake occurrence will be greatly reduced.

It is right to say that randomness comes from symmetry,but it is better to say that randomness comes from unknown asymmetry. When a coin is thrown and if the dice, the throw, the ground and the wind are all perfectly symmetric, the coin will stand on its edge. Only because the throw and the wind are usually not so symmetric and change with each throw in an unknown way, the upper side is front or back at random; if they are known by some accurate measurements, the results can then be predicted with much less random error.

It is quite common in daily life and in technical works to be satisfied with a probable answer either because of incomplete knowledge or because of economy or simplicity, but not to spend much money and efforts trying to predict the exact result of each trial. Theory of probability is the right tool to provide an answer in these cases. Weather prediction is such an example. The error of a prediction of the weather tomorrow or next month was very large 50 years ago; this random error is greatly reduced because of the fast development of the modern sciences and technology and it is sure that this error will be reduced continuously.

Fuzzy ideas come from subjective and qualitative classification, which ignores some factor or factors. "Young man" ignores the age; strong earthquake ignores the magnitude. The necessity of introducing fuzzy terms is simplicity, because of ignorance of the details or unwilling to go to details.

All three types of uncertianty mentioned above exist in SHA. Attenuation of ground motion is always considered to be random. Depiction of the boundary of earthquake province or potential source zone, modeling of earthquake occurrence and attenuation, and evaluation of seismicity parameters such as annual rate, upper limit magnitude (Li,1989) and b value are examples of ignorant or incomplete uncertainty. When expert's opinions are considered,as it is in SHA,fuzzy uncertainty is introduced in experts' opinions. Differences exist not only in the nature of uncertainties but in the methodology of their analysis.

3 UNCERTAINTY ANALYSIS

Lind has expressed the inter–relationship of the three types of uncertainties and pointed out that probabilistic method may be used to analyze random factors, and fuzzy set to analyze fuzzy factors, and multiple subjective probability to analyze ignorant factors (Lind,1985). He pointed also that Bayesian method may be applied to study the relation between random and ignorant factor, but no methods available to study the relations of fuzzy and others.

McGuire has studied jointly the random and the ignorant factors. He emphasized their difference in nature and thus their methods of analysis should also differ . For random factor,say,μ, he

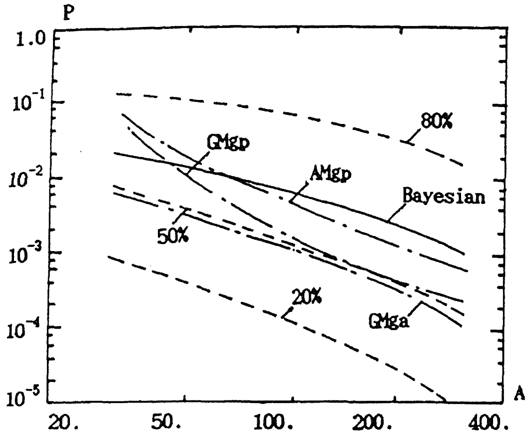

Fig.4 Comparison of various methods of uncertainty analysis

(i = 1,2,...,n), it is

$$P(Y>y) = \Sigma P(Y>y|\mu) P_i \qquad (9)$$

Fig.4 shows the results of various approaches of the effect of uncertainty of ground acceleration attenuation, namely the standard approach of random factors through Eq.1, the fractiles of the Monte Carlo sampling from a log–normal density dirstribution function with the same variance used in the random approach, the arithmatic (Bayesian)or geometric(GMga) mean of the exceedance probability of the Monte Carlo samples for given acceleration, and the arithmatic (AMgp) or geometric (GMap) mean of the acceleration for given probability. The SHA curve for $\mu = 1$ is the result for the average attenuation with no uncertainty; when uncertainty of attenuation is considered, the result is modified to a higher probability of exceedance. It can be seen from a comparison of these results that the result given by Eq.9 from the Monte Carlo samples is the same as that from routine analysis by Eq.1. The modified result is not the one always giving maximum probability of exceedance when compared with all Monte Carlo samples. This is quite reasonable. should be some probability for some sample of factor y of very high value giving higher probability than the one given by using the joint variance.

6 CONCLUSIONS

(1) The sources of random, incomplete and fuzzy factors are actually incompleteness of knowledge of information; for fuzzy factors, details of information, either known or unknown but ignored for simplicity.
(2) Fuzzy and incomplete uncertainties can be measured by subjective probability, which is a measure of the current level of knowledge of the experts.
(3) All three types of uncertainty may be dealt with by Bayesian method.

REFERENCES

Bernreuter,D.L.,Savy,J.B. and Mensing,R.W. 1987, Ground Motion and Engineering Seismology, Developments in Geotechnical Engineering, v.44.

Coppersmith, K.J. and Youngs, R.R. 1986, Capturing Uncertainty in Probabilistic Seismic Hazard Assessments within Interplate Tectonic Envirnments,Proc, 3rd U.S. National Conference on Earthquake Engineering.

Cornell, C.A. 1968,Engineering Seismic Risk Analysis, BSSA, v.58, n.5.

Der Kiureghian, A. and Ang, A.N–S. 1977, A Fault–Rupture Model for Seismic Risk Analysis, BSSA, v.67, n.4.

Hu, Yuxian (ed) 1990, Hybrid Probabilistic Method of Seismic Hazard Assessment, Earthquake Press(in Chinese).

Li, Z.X., Fu, S.Z.and Liu,X.H. 1989, Applicayion of Theory of Fuzzy System on the Evaluation of Upper Limit Magnitude of the Future Earthquakes, Earth– Science in North China, v.7, n.4(in Chinese).

Lind, N.C. 1985,Reconciling Fuzzy Concepts and Probability in Structural Engi– neering, Proc. NSF Workshop in Civil Engineering Application of Fuzzy Sets, .

McGuire, R.K. 1987,Seismic Hazard Uncertainty and Its Effects on Design Earthquake Ground Motions, Proc.International Seminar on Seismic Zonation,Guang– zhou, China.

Savy, J. B.,Berneuter, D. L. and Mensing, R. W. 1986, Overview and Summary of Results for the Project: Seismic Hazard Characterization for the Eastern United States, Proc. 3rd U.S.National Conference on Earthquake Engineering, .

Whitman, R.V. (ed) 1989, Workshop on Ground Motion Parameters for Seismic Hazard Mapping, NCEER–89–0038.

Zhu, Li and Hu,Yuxian 1986, Uncertainty Analysis in Seismic Hazard Assessment, Institute of Engineering Mechanics, State Seismological Bureau, Res, Rept. 86–032 (in Chinese).

Zou, K.Q. andXu, Y. 1989, Fuzzy System and Expert System,Published by Southwest Jiaotong University(in Chinese).