Pergamon

Ann. Nucl. Energy, Vol. 24, No. 10, pp. 751–762, 1997

PII:S0306-4549(96)00052-7
Printed in Great Britain
0306-4549/97 $17.00 + 0.00

ACCOUNTING FOR EXPERT-TO-EXPERT VARIABILITY: A POTENTIAL SOURCE OF BIAS IN PERFORMANCE ASSESSMENTS OF HIGH-LEVEL RADIOACTIVE WASTE REPOSITORIES

E. ZIO† and G. E. APOSTOLAKIS

Department of Nuclear Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, U.S.A.

(Received 19 April 1996; in revised form 28 May 1996)

Abstract—Expert judgments enter several aspects of many scientific endeavors. They are typically employed to interpret data, predict systems' behaviour and assess uncertainties. In particular, expert judgments are expected to be a relevant source of data for use in performance assessments of high-level radioactive waste repositories. In this paper we consider the task of aggregating the judgments provided by experts with the objective of emphasizing a potential source of bias that might affect the results of the analysis. Typically, the analysts combine mathematical and behavioral schemes to obtain the aggregate measures desired for decision-making purposes. Within this approach to data aggregation, mathematical models are used as tools for sensitivity analysis and they should account for between-expert (expert-to-expert) as well as within-expert variability. A practical example regarding a formal expert judgment elicitation exercise for the future climate at Yucca Mountain vicinity is presented. © 1997 Elsevier Science Ltd.

1. INTRODUCTION

Expert judgment is involved in all scientific research mostly in an informal manner, that is its use is implicit and undocumented. Informal judgment is used by a scientist to decide the object of the analysis, to collect the relevant data, and to process and interpret these data. For specific projects requiring inputs from many scientific disciplines, and with outputs that are to be employed in a decision-making process regarding human and environmental safety issues, the informal use of expert judgments may not be satisfactory, as the decisions on these issues must be defendable and the reasons underlying them must be explicitly justifiable. Nuclear reactor safety and reliability issues, as well as the assessments of the performance of repositories for radioactive waste, are typical examples.

†Permanent address: Department of Nuclear Engineering, Polytechnic of Milan, Via Ponzio 34/3, 20133 Milano, Italy.

Hora and Iman (1989), and Keeney and von Winterfeldt (1991), among others, have proposed formal expert elicitation processes devised to gather knowledge about a specific domain from technical experts in a controlled and documented fashion. Typical issues of elicitation include scenario development, model selection and probability encoding. Usually the judgments are elicited in the form of estimates of subjective probability distributions. Probability is interpreted, in this context, as an individual's degree of belief that an event will occur (the event being, for example, that the value of a given parameter lies within a specific range), based on any relevant information available. Further details on the subjective interpretation of probability can be found in Apostolakis (1990).

The elicitation procedure generally involves three parties: the normative experts, the generalists and the substantive experts. The latter are the ones who have deep knowledge and understanding in the domain of interest. The normative experts, on the other hand, have extensive expertise in probability theory and decision analysis practice. The generalists have usually a thorough understanding of the project as a whole, are knowledgeable in the specific issues to be addressed and are aware of the techniques of probability theory and its use in formal expert elicitation processes. In the process of expert elicitation the normative experts elicit judgments from the substantive experts and carefully document their rationale under the supervision of the generalists who provide important inputs from a global view of the process.

Early examples of formal elicitation processes occur in reactor safety studies conducted by the U.S. Nuclear Regulatory Commission, as documented in NUREG-1150 (1987). Many critics of the use of expert judgments have pointed out the need to establish rigorous principles for the collection and interpretation of expert opinions. The methodologies that have been developed with this aim are typically based on a set of steps for the identification of issues, selection of experts, discussion and refinement of issues, training of experts, elicitation, recomposition and aggregation, and documentation. A detailed discussion of these steps is given in Keeney and von Winterfeldt (1991).

When compared with an informal use of expert judgments, formal elicitation is found to increase the scrutability of the judgments and enhance the communication of the results. The training session ensures that the participants are aware of possible cognitive and motivational biases which can affect an expert's judgments. A clear definition of the issues to be addressed eliminates possible misunderstandings and guarantees uniformity of the elicited quantities, thus, in general, improving the accuracy of the assessment.

The major drawbacks to the employment of a formal process of expert elicitation is that it is highly resource and time consuming. Also, compared with informal judgment, formal elicitation may make the expert judgment process somewhat more cumbersome and less flexible. Consequently, the decision to use a formal elicitation process should be carefully considered to ensure that its benefits outweigh its costs.

In this paper we address the problem of aggregating multiple expert judgments with the objective of stressing the existence of a potential source of uncertainty that may significantly bias the results of the analysis. Theoretically, there is no unique, preferred method of aggregating experts' opinions, so that no common convention has been, at present, adopted by the community of practitioners. Many mathematical models exist, but none of them can encompass all of the large variety of situations which can occur when dealing with expert opinions. Typically, the skillful analysts tailor these models to fit them to the specific needs of the study, and combine mathematical and behavioral schemes to obtain the required results. Within this approach to data aggregation, the

available mathematical models offer a valuable tool for sensitivity analysis, in that they allow investigation of the effects that different assumptions (e.g. experts' dependence or independence) have on the combined results. It is therefore important that the mathematical models are properly used, so that the combined results reflect the uncertainties inherent in the individual judgments. This requires that the models account for both within-expert and between-expert variability.

The paper is organized as follows. In the next section, we briefly discuss behavioral and mathematical approaches to the aggregation of expert opinions and point out the source of uncertainty linked to expert-to-expert variability and the possible bias that may arise. Section 3 presents a case study which shows that neglecting this uncertainty may lead to composite distributions which do not give proper credit to the uncertainty expressed by the individual experts. The case study refers to the formal elicitation of future climate in the Yucca Mountain vicinity, presented in a report by DeWispelare *et al.* (1993), within a performance assessment of a radioactive waste repository. The mathematical model used in this case for the aggregation does not preserve the uncertainty actually expressed in the experts' assessments. This is found to be due to the fact that expert-to-expert variability is not accounted for in the analysis. We show that a bayesian approach, which accounts for both within-expert and between-expert variability, leads to very different composite results. In particular, the quantities chosen as examples are the probability of change in precipitation at 7500 years in the future and that of the 10 year average of peak precipitation. The first quantity refers to the difference between the average annual precipitation 7500 years in the future and the present day (1993) average value. The second quantity represents the maximum value of precipitation on the wettest 10 concurrent years (decade) which would be experienced in the Yucca Mountain vicinity in the next 10,000 years, after 1993. The values of this quantity are expressed in terms of average yearly precipitation, where the average is performed over the wettest decade. Finally, in Section 4 we draw some conclusions on the aggregation of expert opinions.

2. AGGREGATING EXPERT JUDGMENTS

Formal elicitation of expert judgments allows the collection of disparate points of view and perspectives from experts with different cultural and educational backgrounds. This should ensure that the issue under analysis receives adequate coverage, unbiased by specific interests or points of view, provided that the range of experts is sufficiently large and comprehensive.

After the elicitation is performed, many individual expert judgments are available. These individual opinions, which may differ substantially due to different backgrounds, assumptions and interpretations, reflect the range of uncertainty regarding the quantity elicited. From a regulatory standpoint, it is important to retain the individual expert judgments so that the regulator can determine the potential impact of different viewpoints, expressed through different aggregation schemes, on the regulatory decision.

On the other hand, the quantity elicited is typically part of a broader analysis and must be used in further calculations. In the case of probabilistic risk assessments for nuclear power plants or performance assessments of high-level radioactive waste repositories, the quantities elicited are typically used as inputs, or parameters, in predictive models of the behavior of the system and its associated phenomena. Therefore, attaining some aggregated

representation that combines the expert opinions on any given elicited issue, or parameter, into a single, coherent representation is often an important step of the process.

Two different approaches exist for the aggregation of expert judgments: the behavioral and the mathematical approach. Extensive reviews of these approaches, and the issues involved, can be found in Genest and Zidek (1986), Cooke (1991), Chhibber *et al.* (1992), Thorne and Williams (1992). The behavioral approach demands that the experts generate a consensus probability distribution. While this may be possible, it certainly brings with it some difficulties related to how the consensus should be reached, what does it represent and to whom does the distribution belong. Structured discussions have been proposed for the achievement of behavioral consensus. Most of these techniques require a facilitator to guide the discussion. In general, the essence of all these structured-discussion techniques is to let all participants express their opinions and to ensure a non-threatening environment.

A somewhat innovative approach to this problem is represented by the Technical Facilitator/Integrator (TFI) approach recently proposed by the Senior Seismic Hazard Analysis Committee (1995). In this approach, it is suggested that a suprapartes technical facilitator team be responsible for the integration of the individual judgments, leading the group of experts to a consensus which properly reflects the available expertise on the issue and the associated uncertainties. Moreover, it is proposed (and here stands much of the innovation) that the TFI itself 'owns' the consensus distribution and is thus directly responsible for responding to any matter concerning it.

In addition to structured discussion, there are non-interactive consensus techniques, such as the Delphi method proposed by Linstone and Turoff (1975), which rely on an iterative cycle of judgmental assessment, communication of the assessments to all participants through a project manager, and reassessment.

In contrast to the behavioral approach in which the experts themselves generate consensus, there exists a large variety of methods for aggregating individual opinions by means of mathematical formulas. Detailed explanations of these techniques can be found in Genest and Zidek (1986), Cooke (1991) and Chhibber *et al.* (1992). A decision-maker, or analyst, is usually responsible for the aggregation of the individual distributions elicited from the experts. Most frequently, in practice, skilled analysts mix and match several aggregation schemes, from the large set available, to address a given problem.

Many of the existing classical models of aggregation present several controversial issues. First of all, it is in general very difficult to quantify possible dependence among experts, although some promising work has recently been published by Clemen and Winkler (1993) and Jouini and Clemen (1994). Moreover, many models of aggregation tend to produce composite distributions which reflect far less uncertainty than that of any of the distributions individually elicited from the experts. As pointed out by Martz (1984), this is due to the fact that between-expert variability is not properly accounted for. A more detailed explanation of this latter point is in order. Suppose that we are given a set of N experts' estimates of a fixed, but unknown, quantity of interest x. Two main components of variability affect the set of estimates: between-expert and within-expert variability. Let, $\tilde{x}_i$, $i = 1, 2, \ldots N$, be the estimate that each expert provides of a fixed, but unknown, quantity, x_i, where x_i is assumed to be a random value of the unknown quantity of interest x from the between-expert variability distribution $g(x_i|x)$. It is further assumed that the unobservable x_i values are statistically independent. The estimate $\tilde{x}_i$ is the observed value of a random variable having distribution $f_i(\tilde{x}_i|x_i)$ which is the subjective probability distribution which represents the within-expert uncertainty, and it is

typically the distribution individually elicited from the experts. The composite distribution is then, by definition, the mixture of two distributions. The first distribution accounts for the between-expert variability and defines a set of expert weights, which, however, are not intended to reflect relative expertise but rather the degree to which each expert's estimates are believed to be a representative sample of the community at large. The second distribution reflects the within-expert uncertainty with regard to the fixed, but unknown, quantity of interest. The aggregated distribution can then be written in the following form:

$$f(x) = \sum_{i=1}^{N} g(x_i|x) f_i(\tilde{x}_i|x_i). \tag{1}$$

Detailed calculations in the case of parametric distributions are illustrated in the work of Martz (1984) and in the report of the Senior Seismic Hazard Analysis Committee (1995), with examples of applications.

There is another viewpoint which is worth mentioning here. If the true value is x and the individual expert's elicitation is $f_i(\tilde{x}_i|x)$, then it is these distributions f_i that consitute the data available to the analyst. The analyst must then derive an a posteriori distribution of his own beliefs conditioned by these data. To do this, he assigns a weight $w_i > 0$ to each expert. These weights should represent the degrees of confidence that the analyst associates with the experts' estimates. The posterior distribution can then be written as

$$\begin{aligned} f(x) &= \sum_{i=1}^{N} w_i f_i(\tilde{x}_i|x), \\ \sum_{i=1}^{N} w_i &= 1. \end{aligned} \tag{2}$$

Computationally, we can say that the analyst is uncertain which expert to believe. He therefore selects one at random, using the distribution of the weights and then selects a value of $\tilde{x}_i$ from $f_i(\tilde{x}_i|x)$). By repeating this process many times, the analyst develops the posterior distribution $f(x)$.

The advantage of this argument is that it avoids reference to the unobservable x_i values. Specifically, it implies that the expert's judgment is dependent upon the real information x and not an unobservable quantity x_i. On the other hand, it leaves the analyst with the difficult and unpopular task of determining and quantifying the *credibility* of the experts.

In the next section we present a case study that shows that the application of the mixture distribution with equal weights can be done in such a way that it fails to fully reproduce the uncertainty actually contained in the individual distributions provided by the experts.

3. A CASE STUDY: EXPERT ELICITATION OF FUTURE CLIMATE IN THE YUCCA MOUNTAIN VICINITY

Expert judgments are expected to be a relevant source of data for use in performance assessments of high-level radioactive waste repositories. In this context, obtaining probability information through elicitation is motivated largely by one fundamental fact: in the

earth and atmospheric sciences, the time and spatial scales involved are so large, and the interrelationships of the phenomena so complex, that obtaining the data to predict reliable estimates of repository performance is not feasible.

As a preparation for an effective review of the U.S. Department of Energy (DOE) safety assessments, the U.S. Nuclear Regulatory Commission (NRC) is evaluating the applicability of expert judgment to the licensing support and regulatory process, including the mechanics of formal expert judgment elicitation to examine the strengths and weaknesses of this method. To aid these evaluations, an actual elicitation was conducted, with regard to the future climate at the Yucca Mountain site in Nevada. The results of this investigation are reported in DeWispelare *et al.* (1993). More specifically, the issue of interest was the definition of potential scenarios of future climate in an area of 50 km radius around the proposed repository for high-level radioactive waste at Yucca Mountain and their associated probabilities of actual occurrence. A panel of five climatologists (who will be indicated here with an alphabetical letter A–E) was selected to provide probability distributions of temperature and precipitation changes at different time epochs. These data will serve as input to the predictive models used in the performance assessments.

Following a familiarization period during which the experts were asked to become acquainted with the problem, a workshop was conducted to discuss the technical issues and train the substantive experts in elicitation procedures. The experts were then allowed to perform their individual research on the issues under analysis. After this period, each expert was individually interviewed to elicit his judgment on the issues. Finally, a group session was conducted to explore aggregation and consensus methods.

The probability distributions provided by the experts were mechanically aggregated by means of a simple weighted average in which each expert's distribution is equally weighted. It is noted that the practice of equally weighting expert judgments has been largely prevalent in most public policy studies, as it eliminates the necessity of making strongly politically charged judgments on the level of expertise of the various experts and it also eliminates the problem of having to assess the value of the weights to be assigned to the various experts.

In the case under examination, the averaging was performed by summing, for a fixed probability value, the different percentiles provided by the experts and then dividing by the number of expert estimates available. For example if $\tilde{x}_i(\alpha)$, $i = 1,2,..5$, is the estimate of the α-th percentile provided by the i-th expert, then the α-th percentile of the composite distribution is given by

$$x(\alpha) = \frac{1}{5}\sum_{i=1}^{5} \tilde{x}_i(\alpha) \tag{3}$$

where, by definition, $P(X < x(\alpha)) = \alpha$.

Figures 1 and 2 report the individual cumulative probability distribution functions (cdfs), provided by the experts, for the change in precipitation (measured in mm) at 7500 years in the future, with respect to the present value (1993), and the 10 year average of peak precipitation (measured in mm/yr) in a decade. The latter quantity is of particular interest to assessing the performance of the repository because it provides some idea of the upper limit of persistent precipitation that is likely to be seen in a relatively short period of

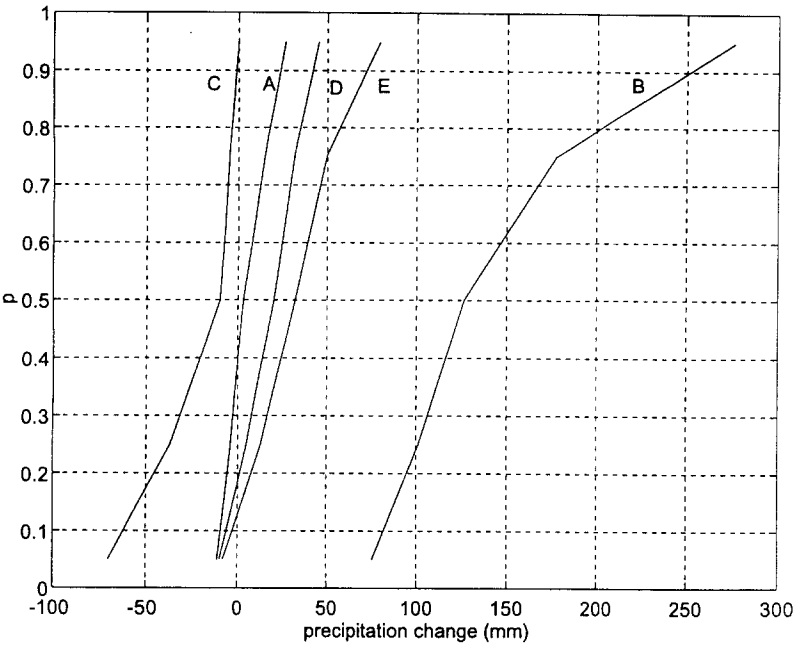

Fig. 1. Cumulative probability distribution functions for the change in precipitation at 7500 years in the future, as provided by the experts.

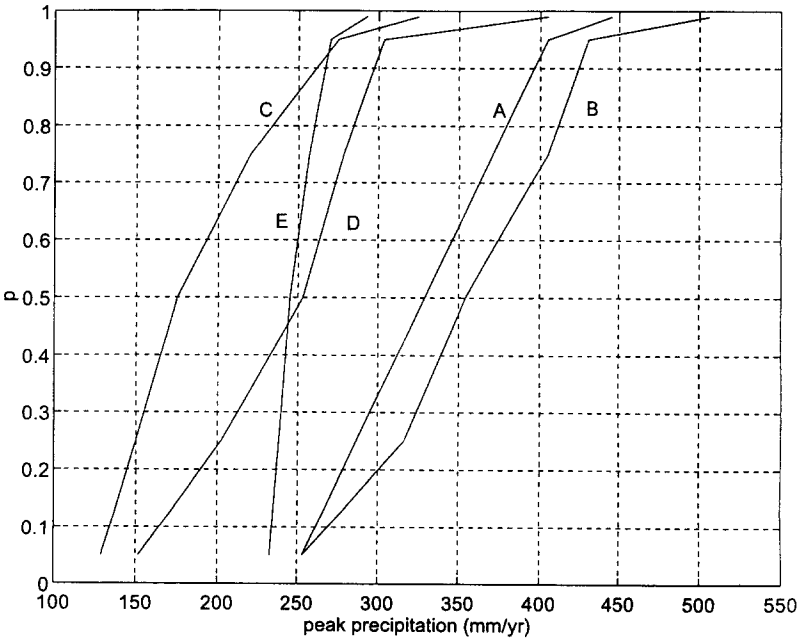

Fig. 2. Cumulative probability distribution functions for the 10 year average of peak precipitation in a decade, as provided by the experts.

time, namely a decade. Tables 1 and 2 contain the elicited point values. The corresponding composite cdfs as obtained by means of the aggregation formula of equation (3) are given by the dashed lines in Figs 3 and 4. The point values for these curves are also reported in the tables. In both cases, the range of potential values of the quantity of interest that is covered by the composite cdfs is narrower than the uncertainty range provided by the set

Table 1. Probability of precipitation change (mm) at + 7500 years

A	B	C	D	E	Average (3)	p
−11.3	75.9	−70	−10	−25	−8.08	0.05
−4.1	101.2	−37.5	4.85	0	12.89	0.25
3	126.5	−10	20	20	31.9	0.5
14.7	177.1	−5	31.36	30	49.63	0.75
26.3	276.3	0	45	50	79.52	0.95

Table 2. Probability of 10 year average of peak precipitation (mm/year) in a decade

A	B	C	D	E	Average (3)	p
253	253.0	129	151.8	233.0	203.96	0.05
286.7	316.3	150	202.4	238.3	238.74	0.25
328.9	354.2	175	253.0	245.0	271.22	0.5
371.1	404.8	220	278.3	256.7	306.18	0.75
404.8	430.1	275	303.6	270.0	336.7	0.95
445.28	506.0	325	404.8	293.0	394.82	0.99

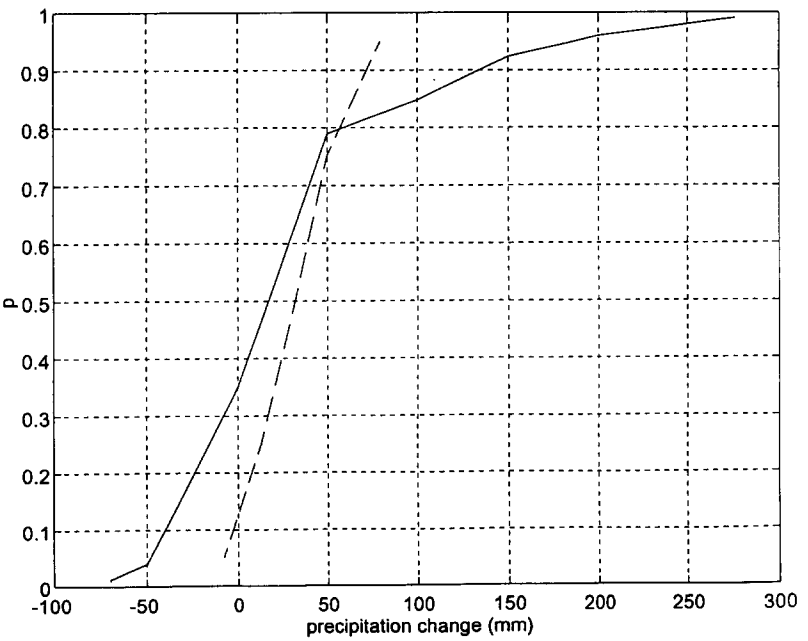

Fig. 3. Aggregated cumulative probability distribution function for the change in precipitation at 7500 years in the future: (– – – –) equation (3); (———) equation (4).

of individual cdfs elicited from the experts. In the case of the 10 year average of peak precipitation in a decade, the range covered by the composite distribution is even less than that of one of the individual cdfs alone, that provided by expert D. This can be seen by direct comparison of Figs 2 and 4 or from the extreme values of the corresponding ranges in Table 2 under the 'average (3)' and 'D' columns, respectively.

A substantially different result is obtained when using the formula of equation (1) for the aggregation. As discussed in the previous section, this way of proceeding allows account to be taken of both the between-expert and the within-expert variability in a theoretically sound fashion. The resulting composite distribution gives a representation of the uncertainty which is coherent with that provided by the set of individually elicited distributions. In the case of equal weights, the between-expert variability distribution $g(x_i|x)$ reduces to a uniform distribution and equation (1) becomes

$$f(x) = \frac{1}{N}\sum_{i=1}^{N} f_i(\tilde{x}_i|x_i). \tag{4}$$

The aggregation is, then, simply performed by summing, for a fixed value of x, the probabilities assigned by the experts to that value. The values obtained with this method of averaging the distributions at fixed values of x are given in Tables 3 and 4. The solid lines in Figs 3 and 4 represent the composite distributions obtained from the application of equation (4). As it can be seen, the uncertainty expressed by the set of individually elicited cdfs is more properly summarized by these composite cdfs than by those obtained using equation (3).

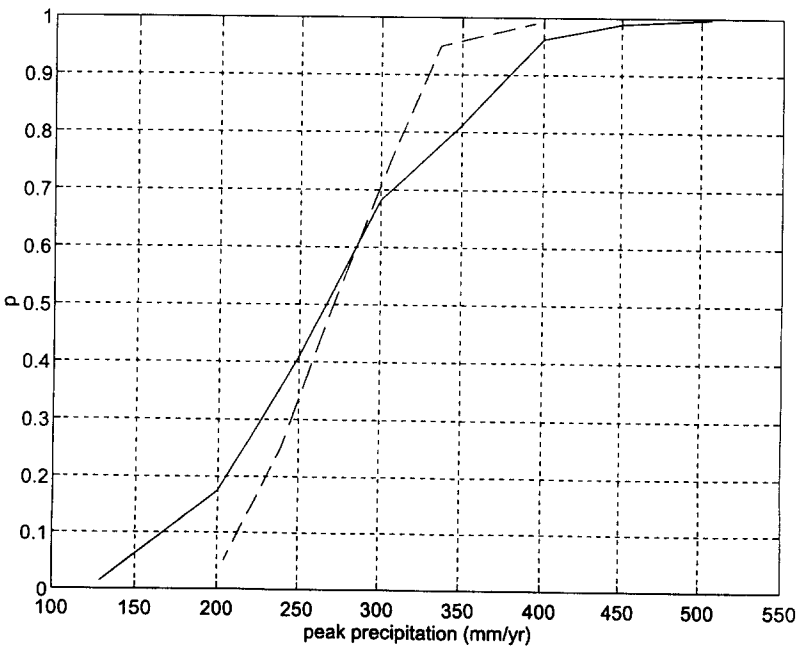

Fig. 4. Aggregated cumulative probability distribution function for the the 10 year average of peak precipitation in a decade: (– – –) equation (3); (———) equation (4).

4. CONCLUSIONS

Expert judgments constitute a relevant source of data for many probabilistic analyses and decision-making processes.

Due to the current state of (incomplete) knowledge in many of the fields involved in modeling for the long term prediction of the performance of a high-level radioactive waste repository, the use of expert judgment in license support activities is necessary and inevitable. The issue of granting a license application has extreme public relevance and all data and analyses contributing to the final decision need to be transparent and unambiguous. The documentation required in support of possible litigation is extensive and can only be adequately supplied by a formal process that emphasizes thorough analysis and complete documentation. These circumstances clearly warrant the use of formal procedures of expert judgment elicitation.

The elicitation approach to expert judgment permits the formal incorporation of subjective expert judgment into the decision process. The information so obtained is intended to supplement other sources, including the results of experiments and observations, and modeling of physical and geochemical processes.

Often, the individual expert judgments obtained through the elicitation process need to be aggregated in a composite form which serves as input to the predictive models used in the assessment.

At present, there is no commonly accepted behavioral or mathematical method of aggregation of opinions. It is, however, widely recognized that it is important to preserve

Table 3. Composite cumulative distribution function for the precipitation change (mm) at +7500 years obtained by means of equation (4)

x	f_A	f_B	f_C	f_D	f_D	Average (4)
−70	0	0	0.05	0	0.005	0.0110
−50	0	0	0.1731	0	0.025	0.0396
0	0.3655	0	0.95	0.1847	0.25	0.3500
50	0.9743	0.0329	1	0.99	0.95	0.7894
100	1	0.2405	1	1	1	0.8481
150	1	0.6161	1	1	1	0.9232
200	1	0.7962	1	1	1	0.9592
250	1	0.897	1	1	1	0.9794
276.3	1	0.95	1	1	1	0.99

Table 4 Composite cumulative distribution function for the 10 year average of peak precipitation (mm/year) obtained by means of equation (4)

x	f_A	f_B	f_C	f_D	f_D	Average (4)
129	0	0	0.05	0.0190	0	0.0138
200	0	0	0.6389	0.2405	0	0.1759
250	0.0470	0.0472	0.8591	0.4852	0.6068	0.4091
300	0.3288	0.1985	0.97	0.9215	0.9912	0.6821
350	0.625	0.4723	0.992	0.9683	1	0.8115
400	0.9215	0.7623	0.996	0.9881	1	0.9624
450	0.9904	0.9605	1	0.9947	1	0.9891
506	0.9958	0.99	1	1	1	0.9972

the individual opinions, as they contain more information than the aggregated distribution, as pointed out in Clemen (1987).

In any case, it is important that the composite distribution obtained be such to account for both between-expert and within-expert variability, thus providing a proper representation of the uncertainty contained in the set of distributions elicited from the individual experts.

In this regard, the case study analysed seems emblematic of a possible bias that may occur. In that study, the judgments of five experts on parameters describing the future climate in the Yucca Mountain region have been combined by means of an equal-weights average. In general, the weighted-average approach (of which the equal-weights is a special case) is simple to apply, provided that the weights have been properly assessed. However, the simplicity can be deceptive. A common, and often overlooked, source of bias arises, in practice, from weighting the fractiles of the distributions or even their moments. This should be avoided as it can lead to gross underestimation of the appropriate amount of uncertainty in the aggregate distributions. The weighting, indeed, should be applied to the distributions directly. As shown in this paper, this way of proceeding has a sound theoretical basis and leads to resulting composite distributions with a broader spread, which more properly accounts for the actual uncertainty in the set of elicited distributions.

As a final observation, we note that more recently the general tendency on the matter of aggregating expert opinions seems to be shifting towards the use of a mixture of behavioral and mathematical techniques, in which the mathematical models of aggregation are considered and evaluated in the context of the overall assessment, thus providing efficient tools for sensitivity analysis.

Acknowledgement—The authors wish to thank the anonymous referees for their useful and insightful comments.

REFERENCES

Apostolakis, G. E. (1990) The concept of probability in safety assessment of technological system. *Science* **250**, 1359–1364.

Chhibber, S., Apostolakis, G. E. and Okrent, D. (1992) A taxonomy of the use of expert judgments in safety studies. *Reliability Engng System Safety* **38**, 27–46.

Clemen, R. T. (1987) Combining overlapping information. *Mgmt Sci.* **33**, 373–380.

Clemen, R. T. and Winkler, R. L. (1993) Aggregating point estimates: a flexible modeling approach. *Mgmt Sci.* **39**, 501–514.

Cooke, R. M. (1991) *Experts in Uncertainty: Expert Opinion and Subjective Probability in Science*. Oxford Univ. Press, New York.

DeWispelare, A. R. *et al.* (1993) Expert elicitation of future climate in the Yucca Mountain vicinity. CNWRA 93-016.

Genest, C. and Zidek, J. V. (1986) Combining probability distributions: a critique and an annotated bibliography. *Stat. Sci.* **1**, 114–148.

Hora, S. C. and Iman, R. L. (1989) Expert opinion in risk analysis: the NUREG-1150 methodology. *Nucl. Sci. Engng.* **102**, 323–331.

Jouini, M. and Clemen, R. T. (1994) Copula models for aggregating expert opinions. *Mgmt Sci.*

Keeney, R. L. and von Winterfeldt, D. (1991) Eliciting probabilities from experts in complex technical problems. *IEEE Trans. Engng. Mgmt* **38**, 191–201.

Linstone, H. A. and Turoff, M. (1975) *The Delphi Method, Techniques and Applications.* Addison-Wesley, Reading, MA.

Martz, H. F. (1984) On broadening failure rate distributions in PRA uncertainty analyses. *Risk Anal* **4**, 15–23.

NUREG-1150 (1987) Reactor risk reference document. Draft copy for comment, Vol. 2, U.S. Nuclear Regulatory Commission.

Senior Seismic Hazard Analysis Committee (1995) An advanced methodology for probabilistic hazard analysis. Draft Report.

Thorne, M. C. and Williams, M. M. R. (1992) A review of expert judgment techniques with reference to nuclear safety. *Prog. Nucl. Energy* **27**, 83–254.