DISCUSSION PAPER

May 2010 ■ RFF DP 10-29

Climate Change Uncertainty Quantification

Lessons Learned from the Joint EU-USNRC Project on Uncertainty Analysis of Probabilistic Accident Consequence Codes

Roger M. Cooke and G.N. Kelly

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org

RESOURCES FOR THE FUTURE

Climate Change Uncertainty Quantification: Lessons Learned from the Joint EU-USNRC Project on Uncertainty Analysis of Probabilistic Accident Consequence Codes

Roger M.Cooke and G.N. Kelly

Abstract

Between 1990 and 2000 the U.S. Nuclear Regulatory Commission and the Commission of the European Communities conducted a joint uncertainty analysis of accident consequences for nuclear power plants. This study remains a benchmark for uncertainty analysis of large models involving high risks with high public visibility, and where substantial uncertainty exists. The study set standards with regard to structured expert judgment, performance assessment, dependence elicitation and modeling and uncertainty propagation of high dimensional distributions with complex dependence. The integrated assessment models for the economic effects of climate change also involve high risks and large uncertainties, and interest in conducting a proper uncertainty analysis is growing. This article reviews the EU-USNRC effort and extracts lessons learned, with a view toward informing a comparable effort for the economic effects of climate change.

Key Words: uncertainty analysis, expert judgment, expert elicitation, probabilistic inversion, dependence modeling, nuclear safety

Contents

Climate Change Uncertainty Quantification: Lessons Learned from the Joint EU-USNRC Project on Uncertainty Analysis of Probabilistic Accident Consequence Codes

Roger M. Cooke and G.N. Kelly*

Introduction

Few subjects are as important, and at the same time as uncertain, as the consequences of climate change. The historical record is ambiguous, the models used to predict climate impacts are uncertain, and the consequences of different outcomes range from the beneficial to the catastrophic. Quantifying uncertainty through a transparent and validated process deserves a high priority. Many of these features, albeit on a smaller scale, characterize risks from nuclear power plants. Considerable effort has been devoted to defensible and transparent quantification of uncertainties in this area. This paper extracts lessons learned for climate change uncertainty quantification from the uncertainty analysis of probabilistic accident consequence codes, COSYMA (EU) and MACCS (USNRC), and provides extensive references..

The joint EU-USNRC uncertainty quantification was initiated and funded jointly by the European Commission and the U.S. Nuclear Regulatory Commission between 1990 and 2000. Although precise cost estimates have not been retrieved, a ballpark estimate for the entire study, including expert remuneration ($15,000 per expert) is US$7 million (2010). The joint study builds on the earlier NUREG-1150 expert judgment exercise (Hora and Iman 1989). There were 2,036 elicitation variables, assessed by 69 experts spread over nine panels. In total, 15,422 individual expert-variable elicitations were performed.

The reports from this study and selected supporting documents are listed in the references. Also included are links and websites from which digitized reports may be downloaded.

* Roger M. Cooke, Resources for the Future, Washington, DC, and Department of Mathematics, T.U. Delft; G.N. Kelly, European Commission, DG Research (ret.), Brussels.

Background and Purpose of EU-USNRC Studies

Accident consequence codes model the adverse consequences of potential accidents in nuclear power plants. Separate codes have been developed with support from the European Commission (COSYMA) and by the U.S. Nuclear Regulatory Commission (MACCS). The scope of these models is depicted in Figure 1.

The objectives of the project were as follows:

- to formulate a generic, state-of-the-art methodology for uncertainty estimation that is capable of finding broad acceptance;
- to apply the methodology to estimate uncertainties associated with the predictions of probabilistic accident consequence codes designed for assessing the risk associated with nuclear power plants; and
- to quantify better and obtain more valid estimates of the uncertainties associated with probabilistic accident consequence codes, thus enabling more informed judgments to be made in the areas of risk comparison and acceptability and therefore to help set priorities for future research.

Figure 1. Scope of Accident Consequence Codes

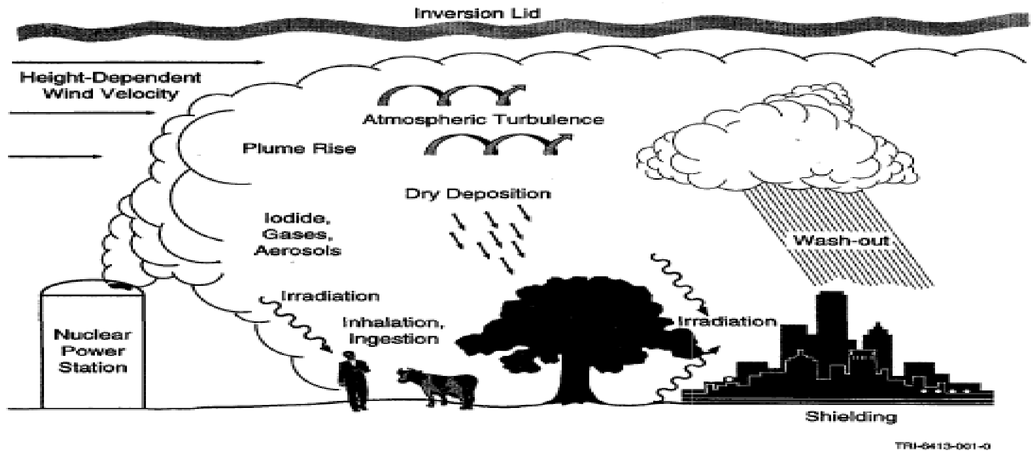
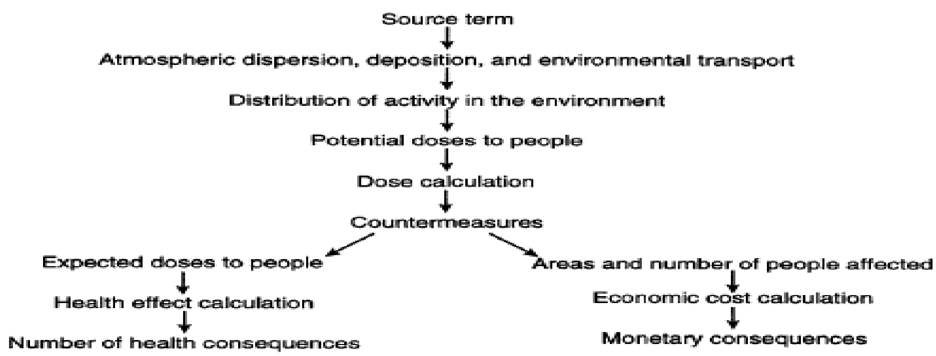


Figure A-1. Dispersion and deposition phenomena considered in an accident consequence analysis.



Uncertainty analyses had been performed with predecessors of both codes, whereby the probability distributions were assigned primarily by the code developers, based largely on literature reviews, rather than by independent experts. Since many input variables, as well as the models themselves, were uncertain, a rigorous and transparent procedure was required to arrive at defensible uncertainty distributions. Both commissions decided to pool their efforts to quantify uncertainty on physical variables and to perform uncertainty analyses on each code separately. The uncertainty quantification was broken into nine separate panels; the number of experts in each panel is shown in Table 1.

Table 1. Expert Panels

Expert panel	*Experts**	*Year*	*Reference*
Atmospheric dispersion	8	1993	Harper et al. 1995 Cooke et al. 1995
Deposition (wet and dry)	8	1993	Harper et al. 1995 Cooke et al. 1995
Behavior of deposited material and its related doses	10	1995	Goossens et al. 1997
Food chain on animal transfer and behavior	7	1995	Brown et al. 1997
Food chain on plant and soil transfer and processes	4	1995	Brown et al. 1997
Internal dosimetry	6	1996	Goossens et al. 1998
Early health effects	7	1996	Haskin et al. 1997
Late health effects	10	1996	Little et al. 1997
Countermeasures	9	2000	Goossens et al. 2001

* The goal was to have half of the experts coming from Europe and the other half coming from the United States. This was not achieved in all panels, for various reasons.

Expert Judgment Methodology

The expert judgment methodology is extensively described in the referenced reports; suffice here to indicate a few principal features:

1. Experts are nominated and selected via a traceable and defensible procedure.
2. Experts undergo a training and familiarization session.
3. Experts prepare their responses prior to the elicitations.
4. Elicitations are conducted individually by a “domain expert” familiar with the subject matter and a “normative expert” experienced in probabilistic assessment.
5. Experts are queried only about the possible results of physical measurements or experiments, and about possible correlations.
6. With a few exceptions, experts also quantify uncertainty with respect to “seed” or “calibration” variables whose true values are or become known within the time frame of the study.
7. Experts write up their assessment rationales, which are published as appendices to the reports.

8. Experts' names and assessments are preserved for peer review, and their names and assessments are published, but names are not associated with assessments in the open literature.

Point (8) is characteristic of most structured expert judgment studies and is designed to discourage "expert shopping," whereby stakeholders or interveners cherry-pick experts to buttress a predefined viewpoint.

Point (6) is designed to enable performance assessment and validation of the resulting combined distributions. Since expert assessments are by their nature subjective, the attempt is made to "calibrate" these assessments against true values of variables from the expert's field. Performance is measured in two dimensions, statistical accuracy and informativeness. Statistical accuracy is measured as the p-value of the hypothesis that the expert's probabilistic statements are accurate in a statistical sense. Informativeness (Shannon relative information) measures the degree to which an expert's distributions are concentrated on a narrow range of possible values. Table 2 shows the number of elicitation questions and number of calibration questions ("seeds") for each panel.

Table 2. Number of Elicitation Variables and Calibration Variables, by Panel

Expert panel	*Questions*	*Seeds*	*Remarks*
Atmospheric dispersion	77	23	
Deposition (wet and dry)	87	19	14 for dry depos. 5 for wet depos.
Behavior of deposited material and its related doses	505	0	No seed questions
Food chain on animal transfer and behavior*	80	8	
Food chain on plant or soil transfer and processes	244	31	
Internal dosimetry	332	55	
Early health effects	489	15	
Late health effects	111	8	*Post hoc* values
Countermeasures**	111	0	Country specific

* Since livestock practices are different in Europe and the United States, the questionnaires were adapted for European and American experts.

** The Countermeasures panel was not part of the joint USNRC-CEC project but was in the EC follow-up project on uncertainty analysis of the COSYMA software package.

The experts' assessments were combined according to two weighting schemes. The equal-weight scheme assigned each expert equal weight, and the performance-based weighting scheme assigned experts a weight based on their performance on calibration variables. Each scheme can be regarded as a virtual expert whose statistical accuracy and informativeness can be assessed in the same way as that of the experts. Table 3 shows the performance of these two weighting schemes.

As a general conclusion, the performance-based decisionmaker exhibits better statistical accuracy and higher informativeness. In most cases, the equal-weight decisionmaker exhibits acceptable statistical accuracy. In one panel (Food chain on soil and plant transfer and processes), the statistical accuracy of both decisionmakers was problematic. This was attributed to the small number of experts (only four) in this panel. For programmatic reasons, primarily to ensure methodological consistency with the earlier NUREG-1150 study[1] that addressed uncertainties in Level 1 and Level 2 Probabilistic Safety Assessment (PSA), the equal-weight decisionmaker was used for the uncertainty analyses, though both decisionmakers are made available, leaving the choice to the discretion of the user.

Table 3. Performance Scores for Equal Weight- and Performance-Based Combinations, by Panel

Case name	*Weighting scheme*	*P-value*	*Mean Information*
Dispersion	Perform	0.9	1.024
	Equal	0.15	0.811
Dry Deposition	Perform	0.52	1.435
	Equal	0.001	1.103
Wet Deposition	Perform	0.25	1.117
	Equal	0.001	0.793
Animal	Perform	0.75	2.697
	Equal	0.55	1.778
Soil and Plant	Perform	0.0001	1.024
	Equal	0.0001	0.973
Internal Dose	Perform	0.85	0.796
	Equal	0.11	0.56
Early Health	Perform	0.23	0.216
	Equal	0.07	0.165
Late Health	Equal	*******	0.28

Point (5) requires that experts assess uncertainty only with regard to observable variables; that is, they do not assess uncertainty on abstract modeling parameters. Indeed, all models are simplifications, and large codes necessarily employ simplified models. The dispersion models in

[1] NUREG 1150 dealt with level 1 and level 2 PSAs, loss of primary systems and loss of containment.

the codes, for example, employ simple Gaussian models with simple schemes for classifying atmospheric stability. More sophisticated models are available but impose a computational burden that does not comport with the computational demands of probabilistic consequence model. Experts are not required to "buy into" the models used in the codes, and indeed, their assessments could be used to quantify models other than those used in the consequence codes.

The restriction to observable query variables means that experts' distributions must be "pulled back" onto the parameter space of the models via a process known as probabilistic inversion. In short, distributions on model parameters must be found, such that when pushed through the models, the results on the observable quantities agree to the extent possible with the (combined) expert distributions. The development of practical techniques for probabilistic inversion was one of the major achievements of this research project.

Lessons Learned

The joint EU-USNRC uncertainty studies represent a benchmark in each of the submodeling areas addressed as well as in quantifying the uncertainty in the risk of nuclear power plants. Moreover, with their reliance on observable query variables, they represent a major methodological advance in the use of structured expert judgment. The major lessons learned are elaborated below.

1. Value of Structured Expert Judgment Process

Structured expert judgment treats the entire uncertainty quantification process as a scientific data collection activity. The value of following a structured and transparent process, as opposed to "best guesses" and "engineering judgment," is very large.

One benefit resides in clarifying the operational meaning of the variables whose uncertainty is quantified. When modelers adopt values adapted from published literature, it may happen that the operational meaning of the published error bars is not the meaning required in the uncertainty analysis. Thus, error bars for dispersion coefficients reported from tracer experiments reflect the variability in the estimates if the entire measurement procedure were often repeated. This is comparable to fluctuations of sample means under repeated experiments. However, the consequence codes are not concerned with an "average" accident, and the target uncertainty concerns downwind concentrations following a *single* release event. The difference between these two uncertainties is considerable (see Kurowicka and Cooke 2006). Figures 2 and 3 show the 5, 50, and 95 percentiles of the lateral spread of a plume in stability conditions D. Figure 2 illustrates the distributions developed by the modelers at the National Radiological

Protection Board (NRPB) and KernForschungszentrum Karlsruhe (KFK). Figure 3 is the result of structured expert judgment.

A second major benefit in structured expert judgment lies in capturing the experts' reasoning in published rationales. These not only illustrate the different thought processes underlying the uncertainty assessments, they also provide valuable introductions to the modeling issues.

Figure 2. 5%, 50%, and 95% Plume Widths (Stability D) Computed by NRPB and KFK

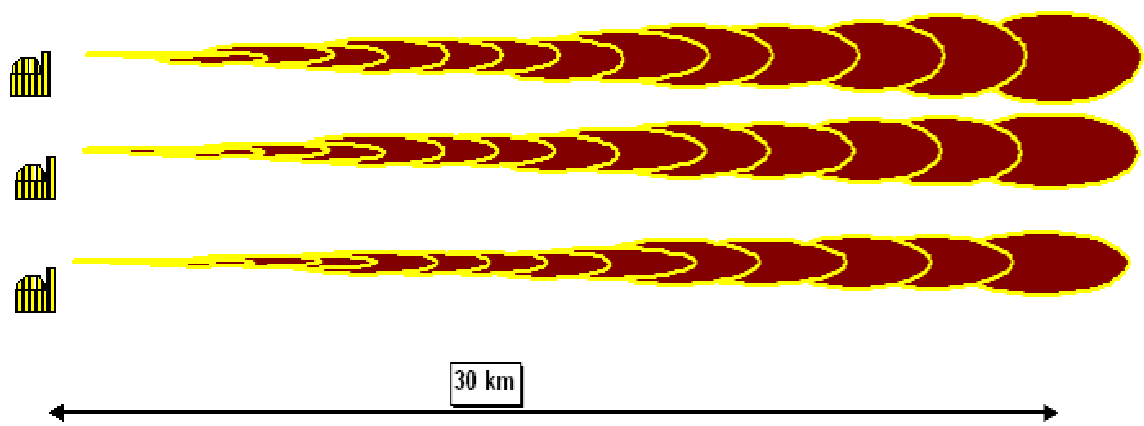
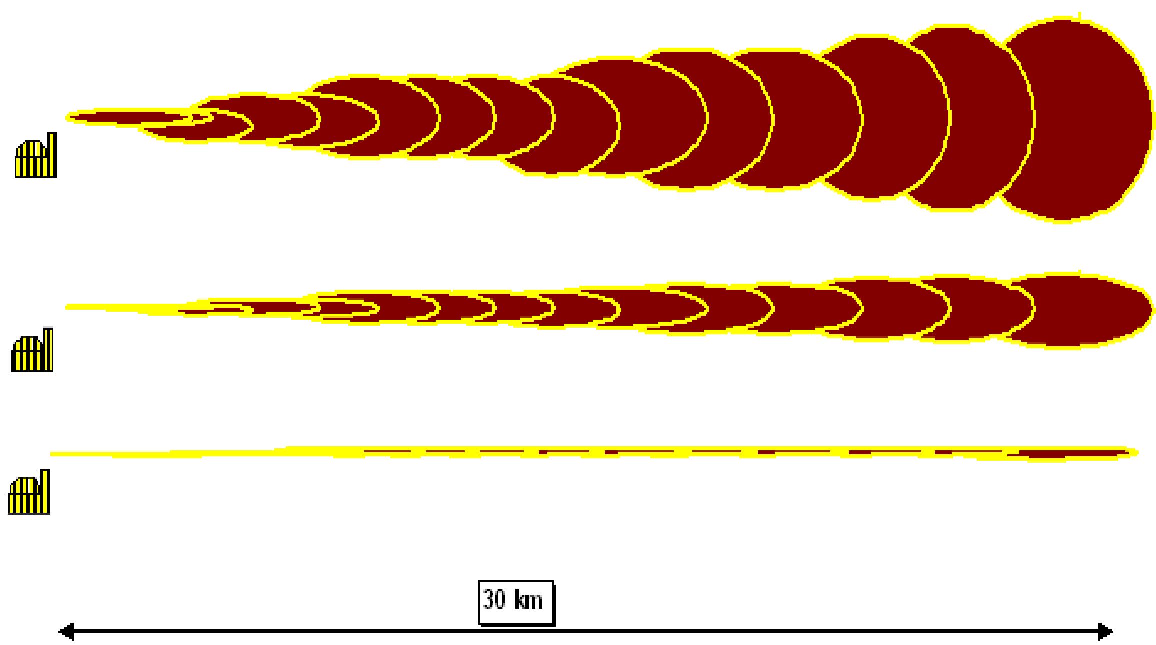


Figure 3. 5%, 50%, and 95% Plume Widths (Stability D) Computed by EU-USNRC Uncertainty Analysis



2. Dependence Modeling and Dependence Elicitation

Prior to the joint study, the question of probabilistic dependence was largely ignored, as if all important dependences were captured in functional relationships, and as if dependence between random variables could be ignored. Of course, this is not remotely true, as best illustrated with a few examples from the joint study:

- the uncertainties in effectiveness of supportive treatment for high radiation exposure in people over 40 and people under 40;
- the amount of radioactivity after one month in the muscle of beef and dairy cattle; and
- the transport of radionuclides through different soil types.

The joint study had to break new ground in dependence modeling and dependence elicitation, and these subjects are treated extensively in the documentation. The format for eliciting dependence was to ask about joint exceedence probabilities: “Supposing the effectiveness of supportive treatment in people over 40 was observed to be above the median value, what is the probability that the effectiveness of supportive treatment in people under 40 would also be above its median value?” Experts quickly became familiar with this format. Dependent bivariate distributions were found by taking the minimally informative copula that reproduced these exceedence probabilities and linking them together in a Markov tree structure. Further developments, generalizing both the choice of copula and the tree dependence structure, are found in Kurowicka and Cooke (2006).

3. Validation of Probabilistic Assessments

Calibration or seed variables were used to assess expert and combined expert performance as probabilistic assessors. Considerable effort went into finding appropriate calibration variables. This has the multiple benefit of raising awareness that subjective probabilities are amenable to objective empirical control, and enhancing credibility in the combined assessments.

4. Combination Methods

Some practitioners believe that expert probabilities should not be combined but simply be presented as multiple views. There is no doubt that individual expert assessments should be part of the published record. However, it is quite unthinkable that all possible combinations of experts in diverse panels should be carried forward to constitute the overall output. Referring to Table 1, this would mean that more than 67 million possible combinations of experts in the various panels would be carried through the entire uncertainty analysis and presented to the user.

There is no practical alternative but to combine the experts in each panel. The equal-weight alternative was used in the NUREG-1150 studies, but the notion that this is the only way of combining experts was criticized. Indeed, equal weighting tends to produce distributions that are significantly more diffuse than any of those provided by experts. Having a well-founded alternative to equal weighting deflects such criticism.

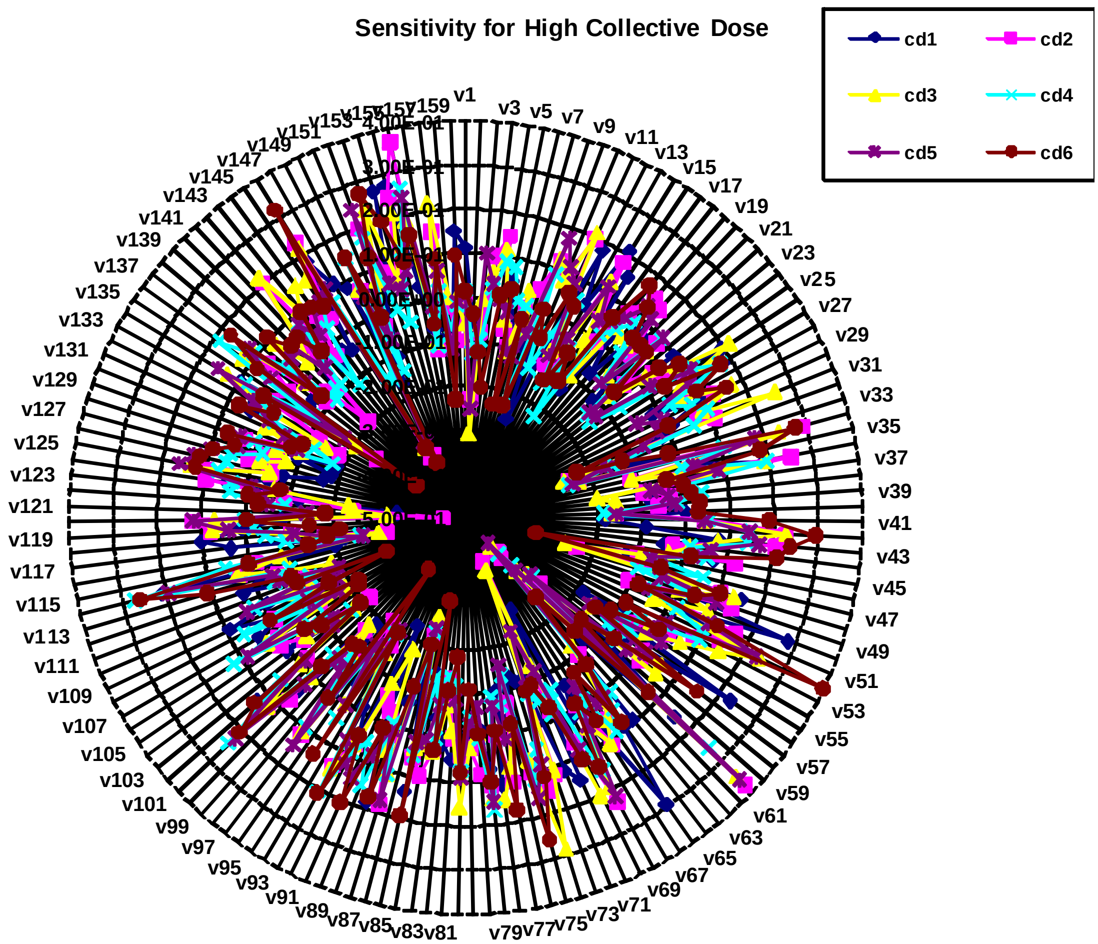
5. Probabilistic Inversion

Essential to the success of the joint study was the decision to query experts only about outcomes of possible measurements. This relieved experts of the burden of assessing parameters of models to which they might not subscribe. Instead, observable quantities were identified that were predicted by the models, and experts were queried on these. New mathematical techniques were developed to pull these distributions back onto the parameter space of the models. These techniques have been further developed since the joint study (Kraan 2002; Kraan and Bedford 2005; Kurowicka and Cooke 2006), and probabilistic inversion is now faster, more flexible, and better grounded mathematically.

6. Sensitivity Analysis and Communication

Extracting useful information from the wealth of data emerging from a large-scale uncertainty analysis is a daunting task. The joint study explored several methods of gauging the importance of the input variables. At the end of the study, however, there was little time to study optimal ways of packaging this information for users. Subsequent work (van Noortwijk and Cooke 2000) further explored graphical communication methods. Figure 4 shows one such tool, developed to illustrate the potential influence of 161 variables on high values of collective radiation dose to six critical organs. A local sensitivity measure was used to capture the influence of the variables on the high-dose regime. A radar graph represents all this information in one image. In A3 format, this gives all the relationships at once, although in normal A4 format it is rather dense.

Figure 4. Radar Graph for Influence of 161 Variables on High Collective Dose to 6 Organs



7. Volitional Uncertainty

Volitional uncertainty is uncertainty with regard to what people will do. First-person uncertainty is uncertainty with regard to what I will do. This type of uncertainty cannot be represented as subjective probability. Uncertainty with regard to what *other* people will do, however, can be so represented. In the Countermeasures module, the joint study had to confront the issue of volitional uncertainty. The reports propose uncertainty distributions for problem owners, risk managers, and local officials, and indeed some of these people were also experts in the panel. Consistent with the overall starting point, uncertainty regarding what people will do must be queried in a way that is divorced from the question of what the expert himself or herself would do. People's actions in the event of an accident must be treated as unknown but potentially observable values of physical measurements.

References

Reports published as a result of the Joint EC-USNRC Project on Uncertainty Analysis of Probabilistic Accident Consequence Codes (under the Third EC-Framework Programme)

Harper, F.T., L.H.J. Goossens, R.M. Cooke, S.C. Hora, M.L. Young, J. Päsler-Sauer, L.A. Miller, B. Kraan, C. Lui, M.D. McKay, J.C. Helton, and J.A. Jones. Probabilistic accident consequence uncertainty study: Dispersion and deposition uncertainty assessment Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities NUREG/CR-6244, EUR 15855 EN, SAND94-1453. Washington/USA and Brussels-Luxembourg, November 1994, published January 1995. Volume I: Main report, Volume II: Appendices A and B, Volume III: Appendices C, D, E, F, G, H.

Cooke, R.M., L.H.J. Goossens and B.C.P. Kraan. Methods for CEC\USNRC accident consequence uncertainty analysis of dispersion and deposition - Performance based aggregating of expert judgements and PARFUM method for capturing modeling uncertainty. Prepared for the Commission of European Communities, EUR 15856, Brussels-Luxembourg, June 1994, published 1995.

Brown, J., L.H.J. Goossens, F.T. Harper, B.C.P. Kraan, F.E. Haskin, M.L. Abbott, R.M. Cooke, M.L. Young, J.A. Jones S.C. Hora, A. Rood, and J. Randall. Probabilistic accident consequence uncertainty study: Food chain uncertainty assessment. Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities NUREG/CR-6523, EUR 16771, SAND97-0335. Washington/USA, and Brussels-Luxembourg, March 1997, published June 1997. Volume 1: Main report, Volume 2: Appendices.

Goossens, L.H.J., J. Boardman, F.T. Harper, B.C.P. Kraan, R.M. Cooke, M.L. Young, J.A. Jones, and S.C. Hora. Probabilistic accident consequence uncertainty study: Uncertainty assessment for deposited material and external doses. Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities, NUREG/CR-6526, EUR 16772, SAND97-2323. Washington/USA and Brussels-Luxembourg, September 1997, published December 1997. Volume 1: Main report, Volume 2: Appendices.

Haskin, F.E., F.T. Harper, L.H.J. Goossens, B.C.P. Kraan, J.B. Grupa, and J. Randall. Probabilistic accident consequence uncertainty study: Early health effects uncertainty assessment. Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities NUREG/CR-6545, EUR 16775, SAND97-2689. Washington/USA and Brussels-Luxembourg, November 1997, published December 1997. Volume 1: Main report, Volume 2: Appendices.

Little, M., C.M. Muirhead, L.H.J. Goossens, F.T. Harper, B.C.P. Kraan, R.M. Cooke, and S.C. Hora. Probabilistic accident consequence uncertainty study: Late health effects uncertainty assessment. Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities, NUREG/CR-6555, EUR 16774, SAND97-2322. Washington/USA and Brussels-Luxembourg, September 1997, published December 1997. Volume 1: Main report, Volume 2: Appendices.

Goossens, L.H.J., J.D. Harrison, F.T. Harper, B.C.P. Kraan, R.M. Cooke and S.C. Hora Probabilistic accident consequence uncertainty study: Uncertainty assessment for internal dosimetry. Prepared for U.S. Nuclear Regulatory Commission and Commission of European Communities, NUREG/CR-6571, EUR 16773, SAND98-0119. Washington/USA and Brussels-Luxembourg, February 1998, published April 1998. Volume 1: Main report, Volume 2: Appendices

Reports published on the Project Uncertainty Analysis of the Probabilistic Accident Consequence Code COSYMA Using Expert Judgement (under the Fourth EC-Framework Programme)

R.M. Cooke, R.M., L.H.J. Goossens, and B.C.P. Kraan. Probabilistic Accident Consequence Uncertainty Assessment Procedures Guide Using Expert Judgement. EUR 18820EN. European Commission, Luxembourg, 2000, Euratom.

Goossens, L.H.J., J.A. Jones, J. Ehrhardt, and B.C.P. Kraan. Probabilistic Accident Consequence Uncertainty Assessment Countermeasures Uncertainty Assessment. EUR 18821EN. European Commission, Luxembourg, 2001, Euratom.

Jones, J.A., J. Ehrhardt, F. Fischer, I. Hasemann, L.H.J. Goossens, B.C.P. Kraan, and R.M. Cooke. Probabilistic Accident Consequence Uncertainty Assessment Using COSYMA Uncertainty from the Atmospheric Dispersion and Deposition Module. EUR 18822EN. European Commission, Luxembourg, 2001, Euratom.

Jones, J.A., J. Brown, F. Fischer, I. Hasemann, L.H.J. Goossens, B.C.P. Kraan, and R.M. Cooke. Probabilistic Accident Consequence Uncertainty Assessment Using COSYMA Uncertainty from the Food Chain Module. EUR 18823EN. European Commission, Luxembourg, 2001, Euratom.

Jones, J.A., F. Fischer, I. Hasemann, L.H.J. Goossens, B.C.P. Kraan, and R.M. Cooke. Probabilistic Accident Consequence Uncertainty Assessment Using COSYMA Uncertainty from the Health Effects Module. EUR 18824EN. European Commission, Luxembourg, 2001, Euratom.