Appearing in Climatic Change DOI: 10.1007/s10584-012-0634-y

Supplementary Online Material

Uncertainty Analysis Comes to Integrated Assessment Models for Climate Change…and Conversely

Roger M. Cooke
Resources for the Future and
Dept. Mathematics, Delft University of Technology
Cooke@rff.org
Feb 24, 2012

This supplementary online material provides background on the following subjects, which are hyperlinked to the corresponding sections of this document.

1. *Structured Expert Judgment in the Joint Study*

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 US Nuclear Regulatory Commission (MACCS).
The scope of these models is depicted in Figure 1.

The objectives of the Joint Study were formulated as

1. to formulate a generic, state-of-the-art methodology for uncertainty estimation which is capable of finding broad acceptance;
2. 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
3. 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.

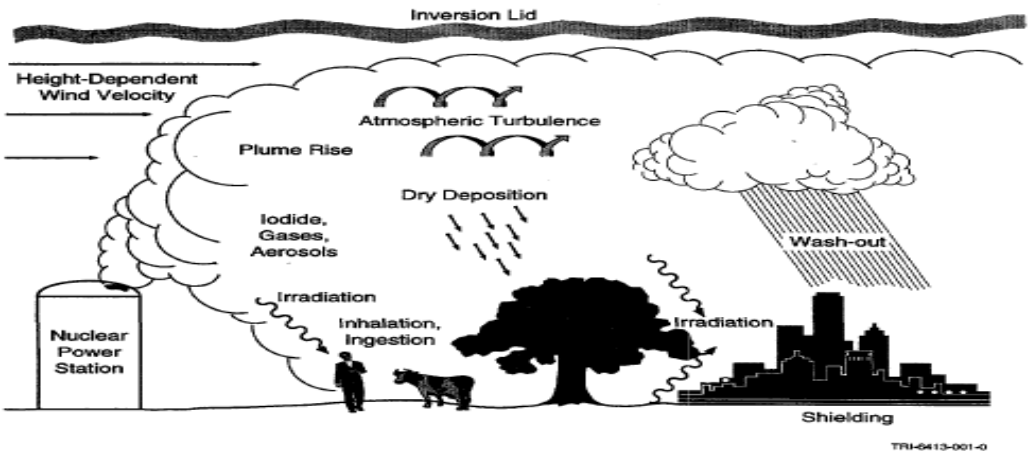

Figure A-1. Dispersion and deposition phenomena considered in an accident consequence analysis.

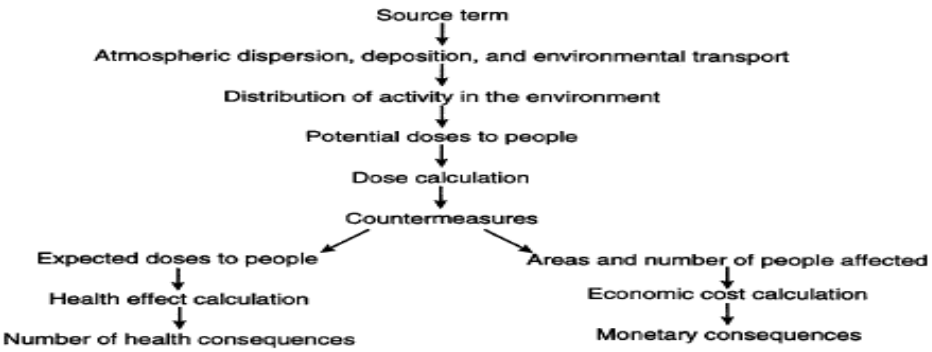

Figure 1. *Scope of Accident Consequence Codes*

Uncertainty analyses had been preformed 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 expert panels. 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. These reports may be downloaded from http://cordis.europa.eu/fp5-euratom/src/lib_docs.htm or http://dutiosc.twi.tudelft.nl/~risk/index.php?option=com_docman&task=cat_view&gid=89&Itemid=13&limitstart=15.

The uncertainty quantification was broken into nine separate panels; the number of experts in each panel is shown in Table 1.

Expert panel	**Number of experts**[1]	**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
Behaviour of deposited material and its related doses	10	1995	Goossens *et al* 1997
Foodchain on animal transfer and behaviour	7	1995	Brown *et al* 1997
Foodchain on plant/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

Table 1: *Expert Panels for Joint Study. Not all experts in each panel assessed all variables, see Table 4.*

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 / 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 from their field whose true values are or become known within the time frame of the study.
7. Experts write up their assessment rationales and these are published as appendices to the reports.
8. Expert names and assessments are preserved for peer review, names and assessments are published, although 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 might cherry pick experts to buttress a pre-defined viewpoint.

[1] The general goal of the panels was to have half of the experts coming from Europe and the other half coming from the USA. This has not been achieved in all panels for various reasons

Point (5) requires that experts assess uncertainty only with regard to observable variables. This entails that experts do not assess uncertainty on abstract modeling parameters. Indeed, all models are simplifications, and large codes necessarily employ simplified models. The dispersion models in 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 other models than those used in the consequence codes. The restriction to observable query variables entails that experts' distributions must be "pulled back" onto the parameter space of the models, via a process known as probabilistic inversion. The development of practical techniques for probabilistic inversion was one of the major achievements of this research project. Since the joint study, inversion techniques have developed substantially, see section 2 of this SOM.

Point (6) is designed to enable performance assessment and to enable validation of the resulting combined distributions. Since expert assessments are by their nature subjective, the attempt is made to validate these assessments against true values of variables from their field. Significant effort went into the selection of calibration variables. Examples of calibration variables are given in Table 2.

Dispersion	Values of near field horizontal and vertical concentration standard deviations as measured in tracer experiments under various conditions, ratio of centerline concentration to off center concentration
Deposition (wet and dry)	Deposition velocities of selected species under specified meteorological conditions, washout coefficients
Internal dose	Retention rates of plutonium and cesium in various target organs, at various times, for children and adults
Soil transport	Penetrations of cesium to various depths at various times, for different soil types
Animal transport	After 4 mo ingestion period, transfer of cesium to meat of dairy cows and sheep, transfer to milk, biological half life in sheep
Early health effects	Dose to red bone marrow after 100Gy/hr dose rate for various cases

Table 2: *Calibration variables for 7 panels*

Performance is measured in two dimensions, namely, 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. It is the probability at which we would falsely reject the hypothesis that the expert's probability statements are accurate. High values close to 1 are good, values near zero are bad. The traditional rejection threshold is 0.05. However, we are not testing and rejecting expert−hypotheses , but using the language of goodness−of−fit to score statistical accuracy. Informativeness (Shannon relative information) measures the degree to which an expert's distributions are concentrated on a narrow range of possible values. High values are good. Complete mathematical definitions are found in (Cooke and Goossens 2008), for derivations and further explanation, see (Cooke 1991). Table 3 shows the number of elicitation questions and number of calibration questions ("seeds") for each panel.

Expert panel	**Number of questions**	**Number of seeds**	**Remarks**
Atmospheric dispersion	77	23	
Deposition (wet and dry)	87	19	14 for dry depos. 5 for wet depos.
Behaviour of deposited material and external dose	505	0	No seed questions
Foodchain on animal transfer and behaviour[2]	80	8	
Foodchain on plant/soil transfer and processes	244	31	
Internal Dosimetry	332	55	
Early health effects	489	15	
Late health effects	111	8	*Post hoc* values
Countermeasures[3]	111	0	Country specific

Table 2: *Number of elicitation variables and calibration variables for each panel*

[2] The Countermeasures panel was not part of the joint USNRC/CEC Project, but part of the EU follow-up project on Uncertainty Analysis of the COSYMA software package.
3 Since the practices of farming with respect to animals is different in Europe and in the USA, the questionnaires were adapted for European and American experts

The experts' assessments were combined according to two weighting schemes. The "equal weight scheme" assigned each expert equal weight, while 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. Figure 2 shows two calibration variables, with expert assessments and the two combinations. Note the non−ovelapping confidence bands and the diversity of expert assessments. Table 4 shows the performance of the experts and of these two weighting schemes. Of the 40 "expert−hypotheses" tested with calibration variables, only 6 would not be rejected at the 5% significance level. Note however, that the power of the statistical tests, as reflected in the number of calibration variables, varies from panel to panel. Nonetheless the pattern found here is consistent with a larger pool of expert judgment studies. Performance assessment for expert probability assessors is the subject of a special issue of Reliability Engineering and System Safety (2008), in which the results of 45 contracted expert judgment studies are analyzed. Further information and analyses may be found in (Woo, 1999), Kallen and Cooke (2002), Cooke et al (2008), O'Hagan et al (2006), Lin and Bier (2008), Lin and Chen (2008, 2009), Aspinall (2010), Flandoli et al (2011), Hora, (2011), Lin and Huang (2012)). The so called "Classical Model" for combining expert judgments was described in (Cooke 1991) and

software is freely downloadable from http://risk2.ewi.tudelft.nl/oursoftware/6-excalibur. Various wild versions are also in circulation for which user discretion is strongly advised.

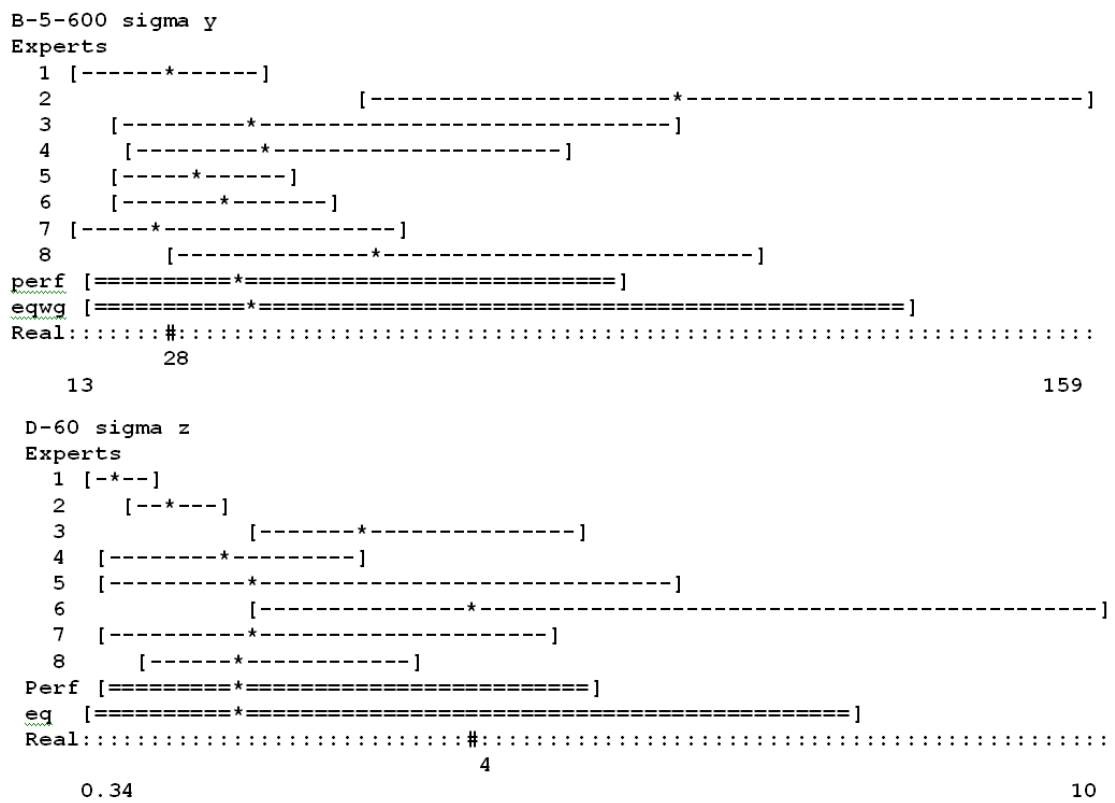

Figure 2 Expert assessments, equal and performance based weighted combinations, and realization for assessments of lateral dispersion (sigma y) in stability conditions B at 600m downwind, and vertical dispersion (sigma z) in stability conditions d at 60m downwind. "*" denotes the median assessment and "[….]" denotes the 90% central confidence band.

As a general conclusion, combining experts enhances statistical accuracy. The performance based combination exceeds the equal weight combination with regard to statistical accuracy and informativeness. In most cases the equal weight decision maker exhibits acceptable statistical accuracy. In one panel (Food chain on soil/plant transfer and processes) the statistical accuracy of both decision makers was problematic. This was attributed to the small number of experts (four) in this panel. For programmatic reasons, primarily to insure methodological consistency with the earlier NUREG-1150 study that addressed uncertainties in Level 1 and Level 2 Probabilistic Safety Assessment, the equal weight decision maker was used for the uncertainty analyses, though both decision makers are made available, leaving the choice to the discretion of the user.

Dispersion				Dry Deposition				Wet Deposition				Animal			
Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns
1	0.0001	2.078	23	1	0.0001	1.953	14	1	0.0001	2.638	19	1	0.001	2.658	8
2	0.0001	1.594	23	2	0.52	1.435	14	2	0.01	1.979	19	2	0.001	2.730	8
3	0.001	1.504	23	3	0.001	1.702	14	3	0.001	1.009	19	3	0.09	1.689	8
4	0.13	1.286	23	4	0.001	1.732	14	4	0.0001	1.028	19	4	0.75	2.697	8
5	0.03	1.092	23	5	0.0001	1.792	14	5	0.001	1.565	19	5	0.01	2.835	6
6	0.005	1.590	23	6	0.001	2.234	14	6	0.0001	1.946	19	6	0.64	2.888	8
7	0.01	1.508	23	7	0.001	1.695	14	7	0.0001	1.252	19	7	0.02	2.821	7
8	0.02	1.840	23	8	0.0005	1.985	14	perf	0.25	1.117	19	perf	0.75	2.697	8
perf	0.9	1.024	23	perf	0.52	1.435	14	eq	0.001	0.793	19	eq	0.55	1.778	8
eq	0.15	0.811	23	eq	0.001	1.103	14								

Soil/Plant				Internal Dose				Early Health				Late Health			
Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns	Exrpt	P-Value	Mean Inf	# Realzns
1	0.0001	2.376	31	1	0.001	1.671	39	1	0.0001	0.834	15	1	-	0.441	0
2	0.0001	1.309	31	2	0.73	0.822	55	2	0.0001	1.375	15	2	-	1.379	0
3	0.0001	1.346	31	3	0.0001	2.003	50	3	0.0001	1.008	15	3	-	1.024	0
4	0.0001	1.607	31	4	0.0001	2.366	39	4	0.0001	0.966	15	4	-	0.507	0
perf	0.0001	1.024	31	5	0.0001	1.205	39	5	0.0001	1.115	15	5	-	0.836	0
eq	0.0001	0.973	31	6	0.005	0.838	28	6	0.0001	0.573	15	6	-	0.599	0
				perf	0.83	0.796	55	7	0.0001	0.410	15	7	-	0.616	0
				eq	0.11	0.560	55	perf	0.23	0.216	15	8	-	0.988	0
								eq	0.07	0.165	15	eq	-	0.280	0

Table 4: *Performance scores for experts, equal weight and performance based combinations, per panel. P–value is the value at which the hypothesis that the expert's probability assessments were accurate would be falsely rejected. Mean Inf denotes the average Shannon relative information with respect to a uniform or log uniform background measure, for all variables (not just the calibration variables). # Realzns is the number of calibration variables used in testing the expert hypotheses. When some experts assessed a subset of the calibration variables, the lowest number of assessed items was used in computing P–values for the entire panel. All computations are with the 2003 version of the program EXCALIBUR. The countermeasure deposited material panels did not employ seed variables. The late health panel queried experts regarding forthcoming results of the Hiroshima and Nagasaki cohort studies. However, due to a change in reporting protocol, true values for the queried variables could not be recovered and the p–values could not be ascertained.*

To compare the results of the Joint Study with previous in house uncertainty assessments Table 5 shows the ratios of the 95th to the 5th percentiles of in-house-expert-modelers at KernForschungszentrum Karlsruhe (KFK) and the same ratios as derived from structured expert elicitation in the Joint Study with equal weights (Harper et al 1995).

Ratio: 95 %-tile / 5%-tiles of uncertainty distributions		
	KfK	**EU-USNRC**

Peak centerline concentration per unit released, 10km downwind, neutral stability	3	174
Crosswind dispersion coefficient, 10 km downwind, neutral stability	1.46	11.7
Dry deposition velocity 1 μm aerosol, wind speed 2 m/s	30.25	300

Table 5: *Ratio of 95- and 5 percentiles of uncertainty distributions computed by the Kernforschungszentrum Karlsruhe (KfK), and by the Joint Study*

2. Probabilistic inversion

Probabilistic inversion as a tool in uncertainty analysis was pioneered in the Joint Study, (Jones et al 2001) although the problem had been encountered earlier in dose response modeling. The power law for dispersion is perhaps the simplest case[3]. Suppose experts give their uncertainty on the standard deviation $\sigma(x)$ of the crosswind concentration following a release, at downwind distances $x = 500m$, $x = 1000$, and $x = 5000m$. Our ability to predict the downwind spreading of a plume is not great, and the uncertainties in $\sigma(x_i)$ $i = 1,2,3$ are considerable. However, these uncertainties are not independent, a plume can never get narrower as it travels downwind. This dependence is captured in the power law $\sigma(x) = Ax^B$, and the uncertainties in σ at *each* downwind distance x should be captured in a single distribution over the coefficients (A,B). Suppose we have a distribution over (A,B) and we draw a large number of samples (a_i,b_i), $i = 1\ldots N$ from this distribution. We want the values $\{a_i 500^{bi} \mid i = 1\ldots N\}$ to mimic the distribution which the experts assigned to $\sigma(500)$, and similarly for $x = 1000$, and $x = 5000$. The operation of finding such a distribution over (A,B) is called probabilistic inversion. The simple power law is believed by no one, as explained in the text. In giving their assessments, the experts may use whatever parametric laws or data resources they like. The analyst is charged to find a distribution over (A,B) which approximately captures the experts' uncertainty. If no such distribution can be found − as sometimes happens − then the modeler is advised that (s)he must revise the model to capture the experts' uncertainty on the model parameters[4].

Such inverse problems are usually hopeless analytically, but suitable numerical approximations can often be found based on variations of the iterative proportional fitting algorithm. The simple problem of the Social Discount Rate $\rho + \eta G_i$, with $G_1 = 1.5$, $G_2 = 2.5$, $G_3 = 3.5$ is described in the text. We choose a diffuse starting distribution for ρ and η. The prescribed distributions for $\rho + \eta G_i$, are $\Gamma(2.5, 1.5)$, $\Gamma(3,2)$ and $\Gamma(5, 2.5)$, where $\Gamma(\mu,\sigma)$ denotes the gamma distribution with mean μ and standard deviation σ. The gamma distributions are not reproduced exactly, rather, we stipulate the 5− 25− 50− 75− and 95−percentiles of the respective

[3] Here follows a more formal definition of probabilistic inversion. Let $X \subset R^M$, and let functions H_i: $X \to R$, $i = 1\ldots n$ be given. Suppose distributions F_i are assigned to H_i, $i = 1\ldots n$. We seek a distribution over X such that the functions H_i take the distributions F_i under this distribution. Such a distribution is a probabilistic inverse of $H_1\ldots H_n$ at $F_1\ldots F_n$. If a probabilistic inverse exists, it is generally not unique and we seek a preferred inverse. If no inverse exists, the problem is infeasible, and we seek a distribution which minimizes some appropriate error function. For more detail, see the literature referenced at the end of this paragraph.

[4] The inversion could be applied to each expert's individual distributions, and then combined across experts. However, the inversion is somewhat labor intensive, and to cut costs it was applied to the combined experts' distributions. This is the distribution used in the uncertainty analysis. Moreover, it may happen that an adequate probabilistic inverse may not exist for some experts, while it does exist for the combination.

gamma distributions. The algorithm starts by drawing a large sample from the diffuse (independent) distributions for ρ and η shown in the left panel of Figure 3.

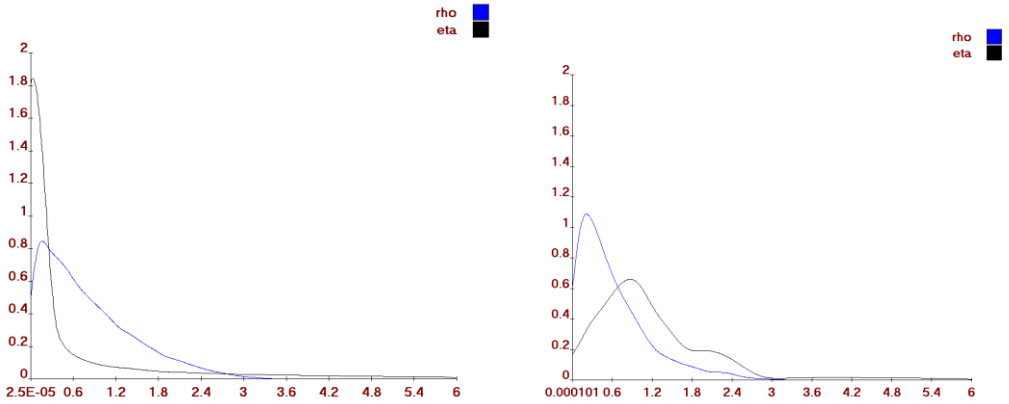

Figure 3: *Marginal distributions of ρ and η before (left) and after (right) probabilistic inversion.*

If *K* samples are drawn, each sample has probability *1/K* in the starting distribution. The iterative proportional fitting algorithm successively reweights these samples so as to comply with each of the 5 stipulated percentiles of each of the three gamma distributions. That is, if we resample the original K samples with probability weights emerging from the algorithm, each of the three functions has percentiles corresponding to the targeted gamma distributions. One cycle of re−weighting thus involves successively fitting 15 constraints. After 100 cycles the cumulative distribution functions in Figure 4 result; g_j are the original gamma distributions, $g_j$pi are the results of probabilistic inversion. One can clearly discern the convergence being forced at the constrained percentiles. Additional constraints and larger starting samples yield better fits. ρ and η are not independent in the resulting distribution; their percentile scatter plot (Figure 4) shows complex dependence. With 15 constraints this problem was feasible.

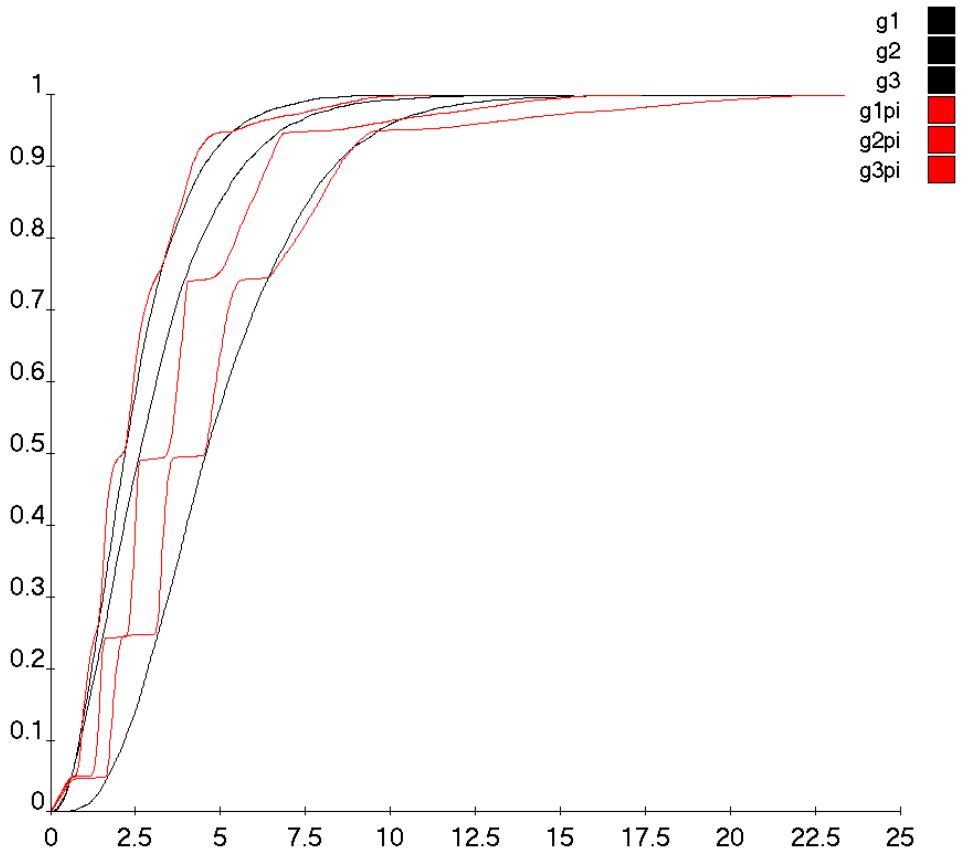

Figure 4: Cumulative distribution functions of g_1,g_2,g_3, and the approximationsg_1−PI,g_2−PI, g_3−PI attained with probabilistic inversion

\

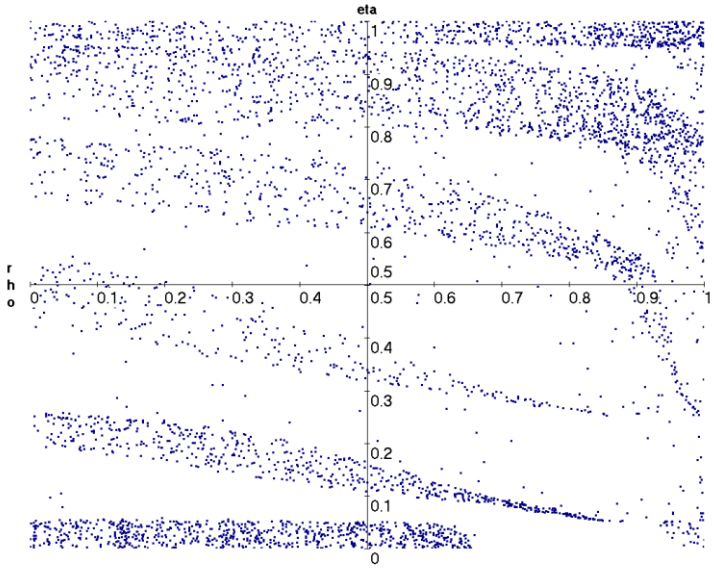

Figure 5*: Percentile scatter plot of ρ and η after probabistic inversion.*

The iterative proportional algorithm was introduced by Kruithof (1937) and re−discovered by Deming and Stephan (1940) and many others. In case of feasibility, Csiszar (1975) showed that the solution is the minimally informative distribution with respect to the starting distribution which satisfies the constraints. In case of infeasibility, little is known about the behavior of iterative proportional fitting. However, variations of this algorithm will always converge to a distribution which minimizes an appropriate error functional (Matus, 2007). For more information on probabilistic inversion see Kraan and Bedford (2005), Du et al (2006) and Kurowicka and Cooke (2006). It may happen that no distribution over model parameters adequately captures the distributions over the target variables and this is an important element in model criticism. Code for performing probabilistic inversion may be freely downloaded at http://risk2.ewi.tudelft.nl/oursoftware/35-universe.

3. Dependence modeling with copulae

This paragraph intends to sensitize the reader to issues that arise in modeling dependence in high dimensional distributions. We assume that a set of marginal distributions is given. If dependence arises through probabilistic inversion, then specialized techniques may be required. We consider the case where dependence information is acquired through expert elicitation.

Although real problems may involve tens or hundreds of variables, we step through a simple example with three exponential variables with unit mean X,Y,Z. Dependence between the variables is anticipated. We must capture this dependence with information elicited from experts, and represent it in a joint distribution for (X,Y,Z). The best known measure of bivariate dependence is the Pearson or product moment correlation. Although this works well for joint

normal variables, it can be problematic for general univariate distributions. The set of feasible product moment correlation matrices for three exponentials is a complex set, and it is unlikely that experts could always give feasible values. For example, the smallest product moment correlation between two exponentials with the same mean is not −1, but about −0.7. A better choice for a bivariate dependence measure is the Spearman or rank correlation. The rank correlation between *X* and *Y* is simply the product moment correlation between the percentiles or quantiles or ranks of *X* and *Y*. All rank correlation values in the interval [−1, 1] are feasible, and they are invariant under monotone increasing transformations of the variables. However, building a high dimensional rank correlation matrix is still challenging[5]. The basic modeling tool is the bivariate copula; that is, a distribution on the unit square with uniform margins. Whatever the joint distribution of *(X,Y)*, if we transform *X* and *Y* to quantiles, the resulting bivariate distribution is a copula. Using graphical techniques illustrated but not explained here, sets of bivariate copulae can be knitted together to enable flexible dependence models of high dimensional distributions (Bedford and Cooke 2002).

Bivariate rank correlations can be elicited from experts in a variety of ways, one of which was indicated in the text. Suppose that in our simple example of three exponentials, each pair *(X,Y), (Y,Z) (X,Z)* is assessed to have rank correlation 0.7. A graphical structure realizing this information is shown in Figure 6. *(X,Y)* and *(Y,Z)* are assigned rank correlation *0.7*. *(X,Z)* are assigned a conditional rank correlation given *Y* of *0.4*; This causes *(X,Z)* also to have rank correlation 0.7. The (conditional) rank correlations may be chosen arbitrarily in the interval [−1,1]. Such a network of bivariate and conditional bivariate constraints is called a regular vine.

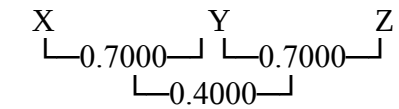

Figure 6*: Regular vine for 3 variables.*

Just as the product moment correlation matrix does not determine the joint distribution, a rank correlation does not determine the copula. Figure 7 shows three copula which each realize rank correlation 0.7.

[5] Every symmetric positive definite matrix with one's on the diagonal is a correlation matrix of some set of random variables. If the set of variables have stipulated univariate distributions, then the set of feasible correlation matrices is strongly constrained. Further, not every correlation matrix is a rank correlation matrix, and not every rank correlation matrix is the rank correlation matrix of a joint normal distribution. The joint normal family realizes every correlation matrix, but its rank correlation matrices are sparse in the set of all rank correlation matrices. With several hundred variables, we always have to deal with partially specified rank correlation matrices, and the issues become more complex. The "completion problem" is the problem of extending a partially specified matrix to a positive definite matrix. Although its exact complexity is not known (Laurent 2001), it is a hard problem. Partially specified structures like that in Figure 5 can be trivially completed by assigning (conditionally) independent copulae to unspecified edges, and this is also the completion with maximum entropy. Such structures can also be easily converted into a sampling algorithm. These facts explain their popularity in uncertainty analysis, for details see (Kurowicka and Cooke 2006).

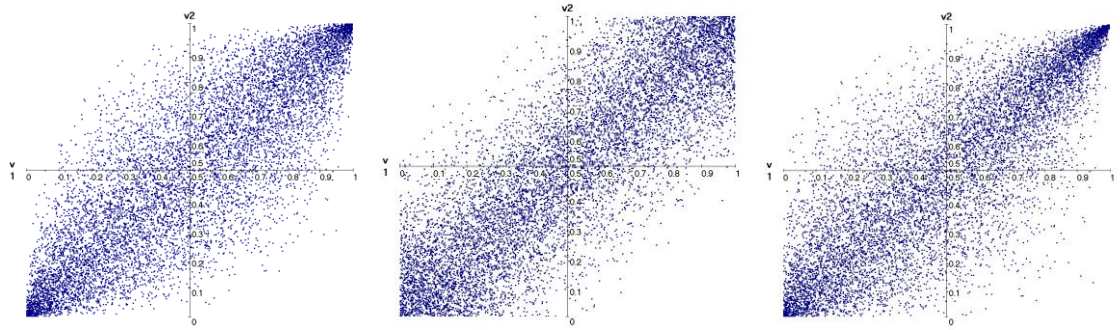

Figure 7: *Three copulae with (rank) correlation 0.7, the normal (left), the minimum information (center) and the Gumbel (right).*

The normal copula is the rank distribution of a bivariate normal distribution. It has the property of tail independence, that is, given a high (low) value of one variable, the conditional probability that the other is high (low) is approximately equal to the independent probability of being high (low). The center copula is minimally informative with respect to the independent distribution, conditional on realizing the rank correlation 0.7. Intuitively, it is the smoothest copula realizing the stipulated correlation. It also has tail independence. Evidently, the normal copula is not minimally informative. The Gumbel copula (right) has upper tail dependence. Each of these copulae, if applied to the structure in Figure 6, realizes the required rank correlation structure. However, if we consider the product $X \times Y \times Z$, the differences as shown in Table 6 are significant.

X × Y × Z **X,Y,Z exponential(1); pair wise rank correlation 0.7**					
Copula	**Mean**	**Variance**	**5% Perc**	**50% Perc**	**95% Perc**
Independent	1.01E+00	7.15E+00	3.15E-03	2.25E-01	4.39E+00
Normal	3.95E+000	2.33E+002	2.83E-004	2.95E-01	1.79E+01
Mininf	2.90E+00	4.40E+01	4.31E-04	2.97E-01	1.43E+01
Gumbel	4.84E+00	4.51E+02	3.69E-04	2.61E-01	2.13E+01

Table 6: *Mean, variance and three percentiles for the product of three exponential variables with unit mean. Except in the independent case, the variables have pair wise rank correlation 0.7 with the normal, the minimal information and the Gumbel copulae.*

As general guidance, if nothing is known about the joint distribution other than the rank correlation, then the minimal information copula is indicated. This was applied extensively in the Joint Study. If tail dependence is suspected, then a tail dependent copula is indicated. Protocols for eliciting tail dependence have not been developed. This is an emerging issue in climate modeling as damage distributions, both in insurance and in finance, exhibit tail dependence. A variety of tools are currently applied in financial mathematics, where the tail independence of the normal copula has fallen in disrepute (Kurowicka and Joe 2011).

4. *Inner and outer measures for climate damages*

The distinction of inner and outer measure is imported from measure theory in mathematics. A strategy for measuring a complex set is to build inner measures and outer measures. Inner measures build simple subsets of the target set and add them up to approximate the target measure from within. Outer measure performs the complementary operation by intersecting simple supersets of the target set. As these simple sets become more refined the inner and outer measures should converge. In mathematics, a set is called "measurable" if the inner and outer measures converge. In the same vein, if measures of climate damages obtained through 'bottom up' and 'top down' methods produce the same results, we may feel more confident in the results. If the subsets used to construct an inner measure are over estimated, the result may of course over estimate the target set. The same applies mutatis mutandis for outer measures.

In measuring the damages from climate change, IAMs have been compiling lists impacted areas; health, crop yield, productivity of labor, storm damage, loss of coastland and migration, etc. For an assessment of the extensive damaging modeling in the IAM FUND see (Smith et al 2009). Every impacted area that we do not think of gets counted as zero. For example, the ocean's phytoplankton accounts for about half of the Earth's oxygen produced by plant life. We have lost about 40% of the ocean's phytoplankton since 1950, and climate change is a prime suspect (Boyce et al 2010). Schlenker and Roberts predict that "Holding current growing regions fixed, area-weighted average yields are predicted to decrease by 30–46% before the end of the century under the slowest (B1) warming scenario and decrease by 63–82% under the most rapid warming scenario (A1FI) under the Hadley III model. " (Schlenker and Roberts, 2009, p 15594). Neither the consequences of loss of phytoplankton biomass nor the recent crop loss predictions are in the IAM's damage models The social institutions essential to economic growth may be impacted by mass migration (other than that due to sea level rise), war, loss of productivity, etc. The difficulty of modeling these features leads easily to their exclusion. Such damage models are inner measures of climate damages.

We can also contemplate an outer measure based on the correlation of economic productivity with temperature. This approach goes back to the following graph of per capita GDP against latitude from Bloom and Sachs (1998).

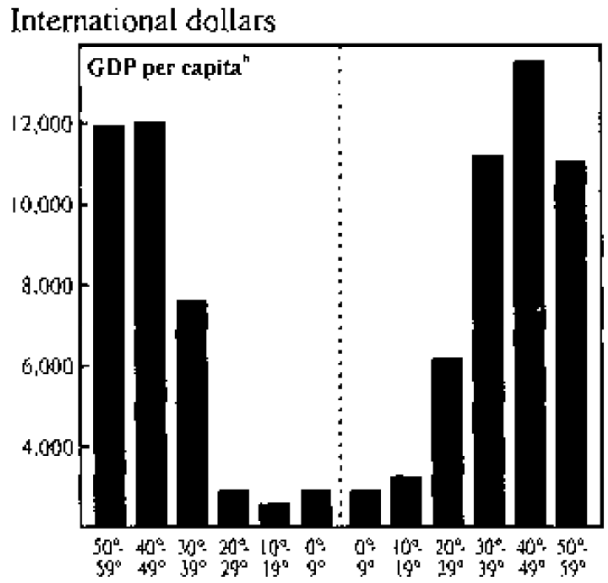

Figure 8: *GDP per capita and latitude (Bloom and Sachs 1998)*

We could build a simple model of the effect of temperature on GDP from this graph. A better starting point is the G-Econ database (Nordhaus 2006, Nordhaus et al 2006) which maps economic output and geographical variables on a $1° \times 1°$ grid. At 45° latitude a grid cell is $45mi^2$ or $68km^2$; the size varies substantially from equator to pole. We consider the non-mineral Gross Cell Product (GCPnm) for 1990 in 1995 USD, converted at market exchange rates. It varies from 0.000103 to 1,155,800 [$\$10^6$] USD(1995). GCP per person (GCPpp) varies from $3.54 to $905,000. There are 27,445 grid cells. Removing grid cells with empty data or zero population and duplications leaves 13,934. A simple regression model can be read from Figure 9. A 1° C rise in average temperature is predicted to result in a 0.0586 decrement on *ln(GCPnmpp).* In other words, if we increase the average temperature in a grid cell by δ, we multiply the *GCPnmpp* by $e^{-0.0586\delta}$. For $\delta = 5C$ this is *0.75*; for $\delta = 10C$ it is *0.56*. In contrast, using the DICE damage function, *5* resp *10℃* warming multiplies output by a factor *0.93* resp. *0.78*.

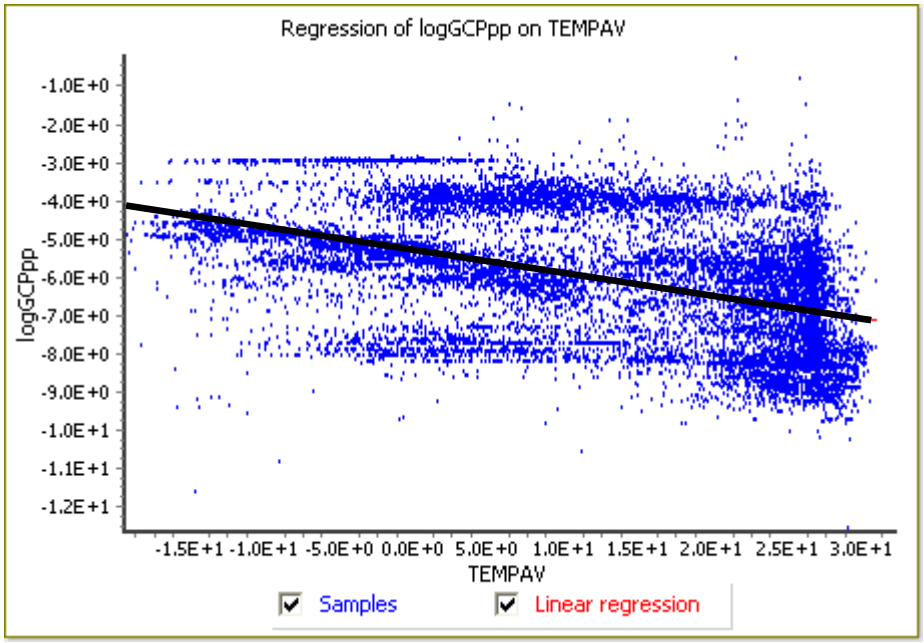

Figure 9: *Regression of ln(GCPnmpp) on average temperature*

Nordhaus et al (2006) propose a model of the form of a production density:

$\ln(y_{i\,j}) = \beta_{0\,j}\,\mathrm{Count}_j + \sum_{k=1..n}\ \beta_k g^k(\mathrm{Geo}_{i\,j\,k}) + \in_{i\,j}$.

y_{ij} is output per km^2 in 1995 international U.S. prices, *i* is the grid cell, *j* is a country or region index, and *k* indexes is the geographical variables. $Count_j$ is country effect, and $\in_{ij}$ is the equation residual. Geographic variables, Geo_{ijk}, are mean annual temperature, mean annual precipitation, mean elevation, ''roughness'' measured as standard deviation of elevation in grid cell, soil categories, and distance from coastline. The g^k are polynomial functions of geographic variables. The Greek variables $\beta_{0\,j}$ are coefficients on regions, whereas the β_k are regression coefficients on geographic variables.

The claim is not that this simple regression is a plausible model; it is a different tack that approaches damages 'from without' rather than from 'from within'. We might discern three different groups in Figure 8. The very high producers with *ln(GCnmpp)* above *−4* seem to prosper through the spectrum of average temperatures. A large middle group is negatively correlated with average temperature. A bottom group seems indifferent to temperature.

The G−Econ data base is a very valuable resource that allows us to approach climate damages in new ways.

5. Social Discount Factor (SDF) and Social Discount Rate (SDR)

Without going all the way back to first principles, c(t) is per capita consumption at time t, whose utility is U(c(t)) and suppose the value of this consumption at time t is valued at t = 0 as

$e^{-\rho t}\,U(c(t))$.

To express this as a factor multiplying c(t), we choose $U(c) = c^{1-\eta} / (1-\eta)$, then $\partial_c U(c) = U'(c) = c^{-\eta}$ and $U(c) = U'(c) \times c / (1-\eta)$. Assuming $U'(c) > 0$, writing $c^{\bullet}(t) = dc(t)/dt$, and using $U''(c)/U'(c) = -\eta/c$, we have

$$U'(c) = \exp(\ln(U'(c)) = \exp(\int (d/dt) \ln(U'(c))\, dt = \exp(\int [U''(c)\, c^{\bullet}(t)/U'(c)]\, dt$$

$$= \exp(-\eta \int [c^{\bullet}(t)/c(t)]\, dt) = \exp(-t\eta G(t))$$

where $G(t) = (1/t) \int_{u=0\ldots t} [c^{\bullet}(u)/c(u)]du$ is the time average growth rate of per capita consumption.

Hence, up to a constant,

$$e^{-\rho t} U(c(t)) = e^{-t(\rho + \eta G(t))} c(t).$$

We denote

$$SDF = e^{-t(\rho + \eta G(t))}; \qquad SDR = \rho + \eta G(t).$$

6. The Bernoulli equation

Let $A(t)K(t)^{\gamma}N(t)^{(1-\gamma)}$ denote output at time t, and put $(d/dt)K(t) = K^{\bullet}(t)$. The model

$$K(t+1) = (1-\delta)K(t) + \phi(t)A(t)K(t)^{\gamma}N(t)^{(1-\gamma)}$$

reduces to a differential equation solved by Jakob Bernoulli in 1695. Write

$$K^{\bullet}(t) + \delta K(t) = B(t)K(t)^{\gamma};\ K(t) > 0;\ B(t) = \phi(t)A(t)N(t)^{(1-\gamma)}.$$

Set $w = K^{(1-\gamma)}$. Then $w^{\bullet} = (1-\gamma)K^{-\gamma} K^{\bullet}$. Dividing by $K(t)^{\gamma}$, this equation becomes:

$$w^{\bullet}(t)/(1-\gamma) + \delta w(t) = B(t).$$

Multiply both sides by $(1-\gamma)e^{(1-\gamma)\delta t}$ to get

$$9e^{(1-\gamma)\delta t} w^{\bullet} + (1-\gamma)e^{(1-\gamma)\delta t}\ \delta w = (d/dt)\,(e^{(1-\gamma)\delta t} w(t)) = (1-\gamma)e^{(1-\gamma)\delta t}\ B(t),$$

The solution is

$$e^{(1-\gamma)\delta t} w(t) = (1-\gamma) \int_{u=o..t}\ B(u)\, e^{(1-\gamma)\delta u}\, du + w(0).$$

Write this as

$$w(t) = (1-\gamma) \int_{x=o..t}\ B(u)\, e^{-(1-\gamma)\delta(t-u)}\, du\ + e^{-(1-\gamma)\delta t} w(0);$$

$$K(t) = [(1 - \gamma) \int_{x=o..t} B(u)\, e^{-(1-\gamma)\delta(t-u)}\, du \; + e^{-(1-\gamma)\delta t}\, K(0)^{(1-\gamma)}]^{1/(1-\gamma)}.$$

If B is constant, this becomes

$$K(t) = [(1 - \gamma)\; B \int_{x=o..t} e^{-(1-\gamma)\delta(t-u)}\, du \; + e^{-(1-\gamma)\delta t}\, K(0)^{(1-\gamma)}]^{1/(1-\gamma)}.$$

Letting $t \to \infty$, we find that the steady state value of capital is given by $(B/\delta)^{1/(1-\gamma)}$.

References

Aspinall, W. (2010) "A route to more tractable expert advice" Nature, vol. 463, 21 January.

Bedford, T.J., & Cooke, R.M., "Vines - a new graphical model for dependent Random Variables"- in Annals of Statistics, vol. 30, No.4, August 2002.

Bloom, D.E. and Sachs, J.D. (1998) "Geography, demography and economic growth in Africa" Brookings Papers on erconomic Activity, vol. 1998, No. 2 207-295.

Brown, J. , Goossens, L.H.J., Harper, F.T. , Kraan, B.C.P. , Haskin, F.E. , Abbott, M.L. , Cooke, R.M. , Young, M.L. , Jones, J.A. , Hora, S.C. , Rood A. , and Randall J., (1997)"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.

Cooke R.M. (1991), Experts in Uncertainty; Opinion and Subjective Probability in Science, Oxford University Press; New York Oxford, 1991.

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

Cooke, R.M. , Goossens, L.H.J., and Kraan, B.C.P. (1995) "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

Cooke, R.M. and Kelly, G.N. (2010) Climate Change Uncertainty Quantification Lessons Learned from the Joint EU–USNRC Project on Uncertainty Analysis of Probabilistic Accident Consequence Codes , RFF DP 10-29.

Cooke, R.M., ElSaadany, S., Huang, X. (2008) On the Performance of Social Network and Likelihood Based Expert Weighting Schemes, Special issue on expert judgment Reliability Engineering & System Safety, 93, 745-756, Available online 12 March 2007, Volume 93, Issue 5, May 2008.

Cooke, R.M., Goossens, L.H.J. (2008) TU Delft Expert Judgment Data Base, Special issue on expert judgment Reliability Engineering & System Safety, 93, 657-674, Nr. 5, May 2008.

Csiszar I. (1975)I-divergence geometry of probability distributions and minimization problems. Ann. of Probab., 3:146–158.

Deming, W.E., and Stephan, F.F. (1940) On a least squares adjustment to sample frequency tables when the expected marginal totals are known, Ann Math. Statist. 40, 11, 427-44.

Du C., Kurowicka D., Cooke R.M.(2006) Techniques for generic probabilistic inversion, Comp. Stat. & Data Analysis (50) 2006, 1164-1187.

Flandoli, F., E. Giorgi, W.P. Aspinall, and A. Neri, (2011) "Comparison of a new expert elicitation model with the classical model, equal weights and single experts, using a cross-validation technique", Reliability Engineering and System Safety 96: 1292-1301.

Goossens, L.H.J. Boardman, J. , Harper, F.T. , Kraan, B.C.P. , Cooke, R.M. , Young, M.L. , Jones, J.A. , and Hora, S.C. (1997) "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

Goossens, L.H.J. , Harrison, J.D. , Harper, F.T. , Kraan, B.C.P. , Cooke, R.M. , and Hora, S.C. (1998) "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

Harper, F.T. . Goossens, L.H.J. . Cooke, R.M. , Hora, S.C. , Young, M.L. , Päsler-Sauer, J., Miller, L.A. , Kraan, B.C.P., Lui, C. , McKay, M.D. , Helton, J.C. , and Jones, J.A. (1995) "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

Haskin, F.E. , Harper, F.T. , Goossens, L.H.J., Kraan, B.C.P. , Grupa J.B. and Randall J. (1997) " 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 Volume 1: Main report, Volume 2: Appendices

Hora, S.C. (2010) "An Analytic Method for Evaluating the Performance of Aggregation Rules for Probability Densities" Operations Research, Articles in Advance, pp. 1–10 issn 0030-364X _ eissn 1526-5463 informs® doi 10.1287/opre.1100.0789

Jones, J..A. , Kraan, B.C.P. , Cooke, R.M. Goossens, L.H.J. Fischer, F. Hasemann, I. (2001) Probabilistic Accident Consequence Uncertainty Assessment Using COSYMA Methodology and Processing Techniques EUR 18827EN European Commission. Luxembourg,Euratom.

Kallen, M.J. and Cooke, R.M. (2002)"Expert aggregation with dependence" Probabilistic Safety Assessment and Management E.J. Bonano, A.L. Camp, M.J. Majors, R.A. Thompson (eds),Elsevier, 2002; 1287-1294. http://risk2.ewi.tudelft.nl/research-and-publications/cat_view/71-publications/84-expert-judgement?start=15

Kraan B.C.P. and Cooke. R.M. (2000) Processing expert judgements in accident consequence modeling. Radiation Protection Dosimetry, 90(3):311-315.

Kraan, B.C.P and Bedford. T.J (2005) Probabilistic inversion of expert judgements in the quantification of model uncertainty. Management Science, 51(6):995 1006, 2005.

Kruithof J..(1937) Telefoonverkeersrekening. De Ingenieur, 52(8):E15–E25.

Kurowicka D. and Cooke R.M..(2006) Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley,

Kurowicka, D. and Joe, H. (2011) Dependence Modeling: Vine Copula Handbook, World Scientific Press, Singapore.

Laurent, M. (2001) Matrix completion problems, available at http://homepages.cwi.nl/~monique/files/opt.ps

Lin, Shi-Woei, and Bier, V.M. (2008) "A Study of Expert Overconfidence" Reliability Engineering & System Safety, 93, 775-777, Available online 12 March 2007. Volume 93, Issue 5.

Lin, Shi-Woei, Cheng, Chih-Hsing (2008) "Can Cooke's Model Sift Out Better Experts and Produce Well-Calibrated Aggregated Probabilities?" Department of Business Administration, Yuan Ze University, Chung-Li, Taiwan Proceedings of the 2008 IEEE IEEM

Lin, Shi-Woei, Cheng, Chih-Hsing (2009) "The reliability of aggregated probability judgments obtained through Cooke's classical model", Journal of Modelling in Management, Vol. 4 Iss: 2, pp.149 – 161

Lin, Shi-Woei, Huang, Ssu-Wei (2012) "Effects of Overconfidence and Dependence on Aggregated Probability Judgments" JMM, appearing March 2012.

Little, M. , Muirhead, C.M. , Goossens, L.H.J. , Harper, F.T. , Kraan, B.C.P. , Cooke, R.M., and Hora S.C. (1997) "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

Matus, F. 2007 On iterated averages of I-projections, Statistiek und Informatik, Universit"at Bielefeld, Bielefeld, Germany, matus@utia.cas.cz.

Meng, C. (2005) "Performance Based Experts Aggregation" Master's thesis, TU Delft, http://risk2.ewi.tudelft.nl/research-and-publications/cat_view/74-master-theses/81-academic-year-2005.

Nordhaus, W.D., (2006), Geography and macroeconomics: New data and new findings, PNAS, March 7, 2006, vol. 103, no. 10.

Nordhaus, W.D., Azam, Q., Corderi, D., Hood, K., Victor,N.M., Mohammed, M., Miltner, A., and Weiss. J., (2006) The G-Econ Database on Gridded Output: Methods and Data1Yale University May 12, 2006.

O'Hagan, A. , Buck, C.E., Daneshkhah, A. Eiser, J.R., Garthwaite, P.H. Jenkinson, D.J. Oakley, J.E. and Rakow, T. (2006) Uncertain Judgements; eliciting expert's probabilities, Wiley, Chichester.

Reliability Engineering & System Safety (2008), 93, 657-674, Available online 12 March 2007, Issue 5, May 2008.

Schlenker, W. and Roberts M.J. (2009) "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change", PNAS September 15, 2009 vol. 106 no. 37, 15594–15598.

Smith, J. , Carney, K., Rodgers, C., Deck, L. and Donahue, J. Tol, R., Ebi, K., and McCarl, B., (2009) "Assessment of Current FUND Modeling, by Sector", Stephanie Waldhoff, U.S. Environmental Protection Agency, 5/1/2009.

Woo, G., (1999). The Mathematics of Natural Catastrophes. Imperial College Press, ISBN 1-86094-182-6.