COMMENTARY:

Messaging climate change uncertainty

Roger M. Cooke

Climate change is full of uncertainty and the messengers of climate science are not getting the uncertainty narrative right. To communicate uncertainty one must first understand it, and then avoid repeating the mistakes of the past.

The recent Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report[1] documents the growth in our understanding of human impacts on the Earth's climate. The inclusion of a chapter on Risk and Uncertainty testifies that large gaps in our knowledge remain. Decisions regarding mitigation and adaptation will impact the Earth we bequeath to our children's grandchildren and these decisions will be taken — or not taken — before these knowledge gaps are closed. Will our descendants praise our foresight or curse our selfishness? Much depends on how climate uncertainty is factored into the current policy debate. There are formidable pitfalls when reasoning under uncertainty, into which both the scientific community and the general population repeatedly fall. This Commentary charts a course.

Table 1 | Word fragment counts presented at the *Uncertainty in Artificial Intelligence* conference in 1985, 2000 and 2012.

	1985	2000	2012
Fuzzy	20%	9%	1%
Belief function	29%	1%	0%
Possibilistic	0%	11%	0%
Certainty factor	20%	0%	0%
Imprecise prob	0%	1%	1%
Random sets	0%	0%	1%
Non-monotronic	5%	0%	0%
Bayes	26%	78%	97%

Current uncertainty narrative

Many influential players in the climate debate do not see uncertainty issues at all. Ranking member of the House Committee on Energy and Commerce of the US House of Representatives John Shimkus told the committee in 2009[2] "The Earth will end only when God declares it's time to be over. Man will not destroy this Earth." US Senator James Inhofe[3] is outraged at "the arrogance of people [who] think that we, human beings, would be able to change what He [God] is doing in the climate". Others, like US presidential candidate Mitt Romney[4], use uncertainty to shift the proof burden: "My view is that we don't know what's causing climate change on this planet, and the idea of spending trillions and trillions of dollars to try to reduce CO_2 emissions is not the right course for us."

Logicians refer to the domain of everyday discourse as the 'natural language', where rules of reasoning are not rigorously defined. The IPCC hoped to raise the debate on climate uncertainty by injecting precisely defined uncertainty qualifiers into the natural language. In 2010 the US National Research Council (NRC)[5] illustrated reasoning under uncertainty about climate change using the calibrated uncertainty language of the IPCC Fourth Assessment Report[6]. The NRC report bases its first summary conclusion on "high confidence" (at least 8 out of 10) or "very high confidence" (at least 9 out of 10) in six (paraphrased) statements[5]:

(1) Earth is warming.
(2) Most of the warming over the last several decades can be attributed to human activities.
(3) Natural climate variability cannot explain or offset the long-term warming trend.
(4) Global warming is closely associated with a broad spectrum of other changes.
(5) Human-induced climate change and its impacts will continue for many decades.
(6) The ultimate magnitude of climate change and the severity of its impacts depend strongly on the actions that human societies take to respond to these risks.

The evident problem with this approach is that the propagation of uncertainty through a chain of inference is conducted in the natural language. Indeed, what is the confidence that all these statements hold? It is not even clear whether "all statements have a 0.8 chance of being true'" means 'each statement has a 0.8 chance of being true' or 'there is a 0.8 chance that all statements are true'. The natural language obscures the gaping difference between these latter two statements. Attempting a rigorous reconstruction of the above chain of inference highlights the limitations of uncertainty propagation in the natural language. Consider the second statement. Does it impute high confidence to 'Earth is warming and humans are responsible', or to the conditional statement 'given that the Earth is warming, humans are responsible'? These are very different statements, and again, the natural language masks this difference. Since the Earth's warming is asserted in the first statement, perhaps the latter, conditional, statement is meant. In that case, the likelihood of both statements holding is the product of their individual likelihoods. If the first two statements enjoy high confidence, then both can hold with only medium confidence ($0.8 \times 0.8 = 0.64$).

The calibrated language translates 'virtually certain' as 99%–100% probability[7].

Suppose the US Nuclear Regulatory Commission licensed nuclear reactors based on the finding each year that each reactor's safety was virtually certain. With 100 commercial nuclear reactors, each with a probability of 0.01 per year of a meltdown… well, do the maths. That is the point: to propagate uncertainty you may have to do some maths. The calibrated language has the important virtue of making problems of uncertainty propagation in the natural language obvious, though apparently not obvious enough.

Back to the past

The lessons of reasoning under uncertainty have been learned many times (see Supplementary Information), but they seem to need re-learning whenever uncertainty erupts in a new field. The artificial intelligence community's experience is illustrative. In 1977 they launched a program to apply their computer chess skills to solving real-world problems, in particular, reasoning under uncertainty in science[8]. Studying the strategies and heuristics of grand masters of science, they concluded that the grand masters did not reason probabilistically, and explored alternative representations of uncertainty, including certainty factors, degrees of possibility, fuzzy sets, belief functions, random sets, imprecise probabilities, and non-monotonic logic, among many others. The NRC reasons as if high confidence in each of their six conclusions were sufficient to convey high confidence in all of them jointly, reflecting the original fuzzy rule for propagating uncertainty.

Proceedings of the premier conference *Uncertainty in Artificial Intelligence* have been digitized since 1985, and provide a unique record of the development of alternative representations of uncertainty. Table 1 shows the relative word fragment count of various approaches. In 1985 the largest component is 'belief function', followed by 'Bayes', 'fuzzy', and 'certainty factor'. Bayes, a proxy for subjective probability, accounts for 26% of the total. By 2000 the balance has shifted; Bayes now accounts for 78% of the count. In 2012 the count is 97% for the word Bayes.

Climate change is the current theatre of alternative uncertainties. A number of economists[9,10] aver that if we don't know the probability distribution, then deep or Knightian uncertainty kicks in, which cannot be characterized by probabilities. As a result, many of the discarded approaches are reappearing. The proponents of deep or Knightian uncertainty perhaps haven't read Knight[11]: "We can also employ the terms 'objective' and 'subjective' probability to designate the risk and uncertainty respectively, as these expressions are already in general use with a signification akin to that proposed." Although the modern foundations of probability, both objective and subjective, date from the publications of Richard von Mises in 1928 and Frank Ramsey in 1931 (see Supplementary Information), Knight was nonetheless able to anticipate in 1921 the objective validation of expert subjective probabilities, which underpins modern science-based uncertainty quantification[12].

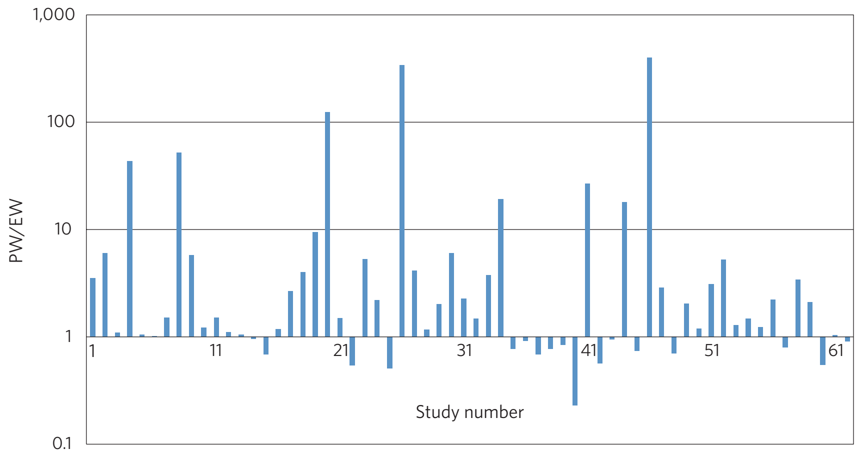

Figure 1 | Performance weight/equal weight (PW/EW) ratios for 62 studies. The ratios concern combined scores for statistical accuracy and informativeness, aggregated over all test/training sets within each study. Data from ref. 22.

Uncertainty quantification

The oft re-learned lesson is this: probability is the logic of partial belief; reasoning under uncertainty must obey the laws of probability. However the probabilities involved are often subjective. Quantitative uncertainty analysis for broad policy questions has always made massive use of expert subjective probabilities[13]. Moss and Schneider advocate this approach for climate uncertainty quantification[14].

The challenge is to render the use of expert subjective probabilities scientific. Early attempts involved systematically eliciting uncertainties from experts and producing a traceable accounting trail. Subsequent developments elaborated and expanded the use of quantitative expert elicitation, involving training in subjective probability assessment, formal elicitation protocols and performance metrics. A series of joint studies by the European Union and the US Nuclear Regulatory Commission employed empirical validation of expert probability assessors, dependence modelling and performance-based differential weighting for combining expert judgments. This "route to more tractable expert advice"[15] underlay a recent application estimating future sea-level rise contributions due to melting of the ice sheets[16], and represents one approach among others to treat expert subjective uncertainty as scientific data. Other approaches[17] that do not employ empirical validation can nonetheless benefit from the large body of performance assessment with structured expert judgment (see the Supplementary Information for extensive detail on expert performance).

Empirical validation

Empirical validation is the hallmark of science. It may come as a surprise that expert subjective probabilities can be, and have been, empirically validated in exactly the way Knight envisaged in 1921. Forty-five professionally contracted studies completed before 2006 were reviewed by Cooke and Goossens[18] from domains including nuclear safety, aerospace and aviation risk, environmental transport, finance, volcanology, banking and public health. In all cases, experts assessed calibration variables from their fields for which the true values were known post hoc.

For example in a study of fine particulate risk[19] experts were asked "On how many days in 2001 did the daily average PM_{10} concentration exceed 50 μg m^{-3} in at least one of the London stations?" As they assess those variables, experts can be treated as statistical hypotheses whose statistical accuracy and informativeness are objectively measured. Performance metrics are used to construct performance-based combinations of expert judgments, also subject to validation. The data from those studies have been made available to researchers and yielded a number of insights, three of which are sketched below (details are in the Supplementary Information).

The first insight concerns experts' overconfidence. Lin and Bier[20] found pervasive overconfidence among experts, measured as the percentage of true values falling outside experts' 90% confidence intervals. However, the differences in expert performance are not random; most expert panels contain statistically accurate experts whose 90% confidence bands tend to contain 90% of the true values. Their results support the case for differential weighting of experts.

Statistical accuracy is only half the story. We want 90% confidence bands that are not only statistically accurate but also informative. The second insight concerns the role of domain expertise and experience in achieving statistical accuracy and informativeness. Using data from the Montserrat Volcano Observatory, Wadge and Aspinall[21] tracked the scores of eighteen specialist volcanologists, and of seven other Earth scientists who act as probabilistic risk assessors. The risk assessors were statistically accurate, but less informative than the most experienced volcanologists. However, some very experienced volcanologists exhibit strong over-confidence.

The third insight concerns performance prediction. Does performance on calibration variables predict performance on the (typically unobservable) variables of interest? When direct observation of the variables of interest is not possible, we rely on expert judgement and need to cross-validate their performance. 'Cross validation' gauges how well performance on a subset of calibration variables (the training set) predicts performance on the complementary subset (the test set). An exhaustive study[22] compares performance-based weighting of experts with equal weighting, for each of 62 studies. Performance weight (PW) combinations of experts based on a training set are applied to a test set and compared with equal weight (EW) combinations. The PW/EW performance ratios for test sets are aggregated over all possible training/test splits for each study. These ratios, shown in Fig. 1, amply attest to the value of performance-based weighting.

The problem of communicating uncertainty cannot be adequately tackled if the communicators don't understand uncertainty. Sprinkling a narrative with uncertainty qualifiers, even if these are given a quantitative interpretation, is not sufficient. Science-based uncertainty quantification is possible, and has been going on for some time in other fields. Much has been learned and climate scientists cannot afford themselves the luxury of repeating the mistakes of the past. ❐

Roger M. Cooke is at Resources for the Future,1616 P st NW, Washington DC, 20036, USA, and Strathclyde Business School, University of Strathclyde, 199 Cathedral Street, Glasgow G4 0QU, UK.
e-mail: cooke@rff.org

References

1. *Intergovernmental Panel on Climate Change Fifth Assessment Report* (Cambridge Univ. Press, 2014); http://www.ipcc.ch/report/ar5/
2. http://go.nature.com/jXapQW
3. http://go.nature.com/gyIzFk
4. http://go.nature.com/vNVp2u
5. National Research Council. *Advancing the Science of Climate Change* 4–5 (The National Academies Press, 2010).
6. *Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties* (IPCC, 2005); http://go.nature.com/fQk55s
7. Mastrandrea, M. D. *et al. Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties* (IPCC, 2010); http://go.nature.com/PvUJbk
8. Feigenbaum, E. A. in *Proc. Int. Joint Conf.on Artificial Intelligence and National Computer Conf.* (1977); http://go.nature.com/VW6q48
9. Henry, C. in *Laboratoire d'Econométrie de l'Ecole Polytechnique Chair Développement Durable Cahier* (Ecole Polytechnique, 2006); http://go.nature.com/FxkZxR
10. Stern, N. *Am. Econ. Rev.* **98,** 1–37 (2008).
11. Knight, F. H. *Risk, Uncertainty, and Profit* III viii.1 (Harper & Row, 1921).
12. Knight, F. H. *Risk, Uncertainty, and Profit* III viii.43 (Harper & Row, 1921).
13. Cooke, R. M. *Climatic Change* **117,** 467–479 (2013).
14. Moss, R. H. & Schneider, S. H. in *Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC* (eds Pachauri, R., Taniguchi, T. & Tanaka, K.) 33–51 (World Meteorological Organization, 2000).
15. Aspinall, W. P. *Nature* **463,** 294–295 (2010).
16. Bamber, J. L. & Aspinall, W. P. *Nature Clim. Change* **3,** 424–427 (2013).
17. Morgan, M. G. *Climatic Change* **108,** 707–721 (2011).
18. Cooke, R. M. & Goossens, L. H. J. *Reliab. Eng. Syst. Safe.* **93** (special issue), 657–674 (2008).
19. Cooke, R. M. *et al. Environ. Sci. Technol.* **41,** 6598–6605 (2007).
20. Lin, S-W & Bier, V. M. *Reliab. Eng. Syst. Safe.* **93,** 711–721 (2008).
21. Wadge, G. & Aspinall, W. P. in *The Eruption of Soufriere Hills Volcano, Montserrat, from 2000 to 2010* (eds Wadge, G., Voight, B. & Robertson, R. E.) Memoir 39, 211–224 (Geological Society London, 2014).
22. Eggstaff, J. W., Mazzuchi, T. A. & Sarkani, S. *Reliab. Eng. Syst. Safe.* **121,** 72–82 (2014).

Additional information

Supplementary information is available in the online version of this paper.

COMMENTARY:

A balanced-efforts approach for climate cooperation

Robert C. Schmidt

Focusing on policies and effort costs rather than emissions may facilitate climate negotiations and improve the chances of reaching a successful agreement. The effort costs of a country comprise investments in low-carbon technologies, in addition to direct mitigation costs.

In the past, climate negotiations have focused primarily on emissions targets. Stiglitz, however, argues that it would be easier to negotiate about taxes[1]. In his view, the advantage of a common tax over the Kyoto approach would be that most of the distributional debate is sidestepped. In particular, under the Kyoto approach, obtaining the right to pollute is like receiving a gift. Hence, countries may struggle for the best 'deal', which can make an agreement difficult to achieve. In an earlier contribution, Schelling[2] suggests that countries choose their own policy instruments when contributing to climate stability. He argues that a proposal should specify policies, such as taxes, regulations, or research and development subsidies,