Climatic Change (2007) 85:19–31
DOI 10.1007/s10584-007-9315-7

Expressions of likelihood and confidence in the IPCC uncertainty assessment process

James S. Risbey · Milind Kandlikar

Received: 22 February 2005 / Accepted: 13 July 2007 / Published online: 6 September 2007

Abstract Communication of uncertainty information in Intergovernmental Panel on Climate Change (IPCC) assessments has evolved through successive reports to provide increasingly formal classifications for subjective and objective information. The first IPCC assessments provided uncertainty information in largely subjective form via linguistic categorizations depicting different levels of confidence. Recent assessments have codified linguistic terms to avoid ambiguity and introduced probabilistic ranges to express likelihoods of events occurring. The adoption of formal schemes to express likelihood and confidence provides more powerful means for analysts to express uncertainty. However, the combination of these two metrics to assess information may engender confusion when low confidence levels are matched with very high/low likelihoods that have implicit high confidence. Part of the difficulty is that the degree to which different quantities in the assessments are known varies tremendously. One solution is to provide likelihood information in a scheme with a range of different precision levels that can be matched to the level of understanding. A version of this scheme is also part of the IPCC uncertainty guidance and is described here.

1 Introduction

Every assessment of climate change is faced with the need to characterize and communicate uncertainties in the state of understanding. This has long been a contentious process. For example, the 'Charney' report (National Research Council

J. S. Risbey (✉)
CSIRO Marine and Atmospheric Research, Hobart, Tasmania 7001, Australia
e-mail: james.risbey@csiro.au

M. Kandlikar
Institute for Asian Research, University of British Columbia, Vancouver, British Columbia V6T1Z2, Canada

1979) characterized climate sensitivity as 3°C ± 1.5°C, sparking a long running discussion of the meaning of the error range. As climate assessments moved into the Intergovernmental Panel on Climate Change (IPCC) process and attracted larger audiences, pressure has mounted to formalize the characterization of uncertainties. The IPCC has provided a forum to develop standard, useful formats for communicating uncertainty. The first IPCC reports highlighted subjective judgements by categorizing results according to various linguistic expressions of confidence: "we are certain of ... ", "we calculate with confidence that ... ", "our judgement is that". Yet, considerable ambiguity remained in the interpretation of these terms and in the lack of any formal method to present quantitative information and likelihoods.

Formal consideration of uncertainty in assessments of the IPCC began with the Third Assessment Report (TAR) (Houghton et al. 2001). As a part of the TAR process, Moss and Schneider (2000) wrote a guidance document that developed a framework for representing and communicating uncertainty in quantitative and qualitative terms. A critical aspect of the Moss and Schneider proposal was a scheme to link qualitative descriptors of uncertainty to quantitative metrics. They proposed a subjective five-unit scale that mapped quantitative ranges of subjective confidence to linguistic descriptors of confidence. They also recognized that in some cases, scientists might want to either supplement or supplant subjective quantitative judgements with descriptions of uncertainty that are qualitative in nature. Accordingly, they also developed the qualitative schema shown in Table 1, which differentiated between the quality of evidence and the level of expert consensus on a particular topic. The Moss and Schneider scale is Bayesian in that it does not make explicit the distinction between subjective and frequentist forms of uncertainty.

This distinction has emerged in representation of uncertainty in the IPCC fourth assessment (AR4) process. Guidance documents on uncertainty communication produced for the AR4 (IPCC 2006) distinguish between *likelihood* and *level of confidence*[1] in representations of uncertainty. In these documents 'likelihood' has a frequentist connotation and refers to a probabilistic assessment of some well-defined outcome having occurred or expected to occur in the future (see Table 2). Subjective expression of uncertainty is introduced in the AR4 via a 'level of confidence' scale. The 'level of confidence' is based on the degree of understanding in the expert community using a probabilistic formulation (see Table 3). In traditional use, likelihoods are based on existing data (from observations or models), and can be applied to all cases where the classical definition of probability based on past counts applies. Levels of confidence can be applied when such data are incomplete and subjective judgement is required.

Separating the likelihood of, and the level of confidence in, a statement has also raised the possibility that likelihood and level of confidence could be combined in making assessments. Indeed, guidance documents for the AR4 assessment of uncertainty state that "it is possible to have high confidence in a finding indicating that climate change would lead to a low probability of some outcome and conversely to have low confidence in a finding that climate change would lead to a high probability of another outcome" (Manning and Petit 2004). While it is possible to combine likelihood and confidence estimates in this way, it is not necessarily

[1]The term 'level of confidence' is distinct from that of a 'confidence interval' as normally used in statistics.

Table 1 This table is reproduced from Moss and Schneider (2000) and provides a qualitative means to express levels of confidence based on the level of expert agreement and the relevant evidence

	Amount of evidence →	
Level of agreement ↑	Established but incomplete	Well established
	Speculative	Competing explanations

The two-by-two matrix allows 'low' or 'high' determinations of consensus and evidence to determine the appropriate confidence expression.

meaningful in all cases. In the following section we explore a variety of possible ways to combine likelihoods and level of confidence. The combinations we consider are using only likelihood, using only confidence levels, and using the two together.

The focus on uncertainty in this paper is on characterizing uncertainty in 'outcomes'. By this we mean outcomes of some event or well-posed question that are subject to quantification or can be expressed in probabilistic form. What is important in this case is that the event and variable be well specified. We are not concerned here about characterizing the event according to uncertainty typologies, only in expressing the outcome. Similarly, we are not concerned here with the assessment of uncertainty in more general knowledge claims which do not have quantifiable metrics. We assess each of the likelihood and confidence schemes below only in terms of their utility for expressing quantifiable outcomes.

2 Likelihood versus levels of confidence: three alternatives

On the face of it, the distinction between likelihood and level of confidence sounds unproblematic and almost innocuous. We argue however, that care must be taken in applying these differing definitions if the goal of effective communication is to be achieved, especially when they are used in combination. In what follows, we examine pros and cons of using likelihood and levels of confidence individually and in combination by defining three alternative options for using likelihood and levels of confidence in communication of uncertainty. We then go on to suggest one alternative for how the two approaches can be fruitfully used in combination.

2.1 Alternative 1: use only likelihood

Alternative 1 would simply use likelihoods and can be interpreted in two ways. The first is the frequentist/classical view and ignores or downplays subjective uncertainty.

Table 2 Likelihood defined as a probabilistic assessment of some well-defined outcome having occurred or occurring in the future

Terminology	Likelihood of occurrence (%)
Virtually certain	> 99
Very likely	> 90
Likely	> 66
About as likely as not	33 to 66
Unlikely	< 33
Very unlikely	< 10
Exceptionally unlikely	<1

Table based on IPCC (2006).

Table 3 Characterizations of levels of confidence expressed in terms of the odds of being correct

Terminology	'Odds' of being correct
Very high confidence	At least 9 out of 10 chance
High confidence	About 8 out of 10 chance
Medium confidence	About 5 out of 10 chance
Low confidence	About 2 out of 10 chance
Very low confidence	Less than 1 out of 10 chance

Table based on IPCC (2006).

On the plus side, this is clear and works well for traditional science problems (Ravetz 1971), especially those based on empirical observations. It also provides scientists with a greater degree of comfort by making a 'clean' separation between objective and subjective knowledge [assuming this were possible (Schrader-Frechette 1984)]. Any information that does not confirm to the norms of classical statistics is discarded in this approach. Unfortunately, this alternative does not always work well for climate change because only a limited set of quantities can be expressed in full likelihood terms. This is effectively the pre-IPCC approach where uncertainty was discussed in as much as rigorously quantifiable measures were available (and avoided if not). Using only classical likelihoods delimits the boundaries of knowledge to a small set of findings based either on the historical record or from models rigorously calibrated to historical data. This makes it difficult to provide meaningful scientific advice to policy makers on a large number of questions where history is unlikely to be a suitable guide for the future.

The second view of likelihood assessments acknowledges that any likelihood assessment would contain subjective elements. With this broader view, one need not limit the domain of applicability of likelihood to problems rich in past count data. However, the more subjective the likelihood assessment, the more the need to evaluate that subjectivity, and the more the assessment would be improved by adding information on confidence levels as well. One way to reduce subjective and uncertain elements of a likelihood assessment is to render these elements conditional—declare assumptions and hold them fixed. For example, one may express likelihoods of change conditional on a particular emissions scenario. This can increase the utility of likelihood statements, though one is still left with a need to assess and communicate the quality of the conditionals. This last step is not very amenable to a likelihood-only scheme.

In summary, if we view likelihoods in strict frequentist terms, they have limited application in addressing climate change issues. If on the other hand we imbue likelihoods with subjective content (or express them conditionally), then they have wider application in climatology, but that subjectivity (conditionality) needs to be evaluated. However, the likelihood scheme itself is inappropriate for subjective evaluations and needs to be supplemented with a qualitative framework.

2.2 Alternative 2: use only levels of confidence

As noted above, it is rare to have rigorous likelihood data for all but a few variables. The likelihood of future events can be formally determined from models, but models have many subjective elements (van der Sluijs et al. 1998; Shackley et al. 1999; Murphy et al. 2004). Thus, it could be argued that a majority of data relevant to assessment of future climate change has embedded subjective elements. The quality of the data, and suitability of models for the questions posed determines the level of confidence (Risbey 2002). The early IPCC reports (Houghton et al. 1990, 1996) took a level of confidence approach without quantification. Instead they used linguistic approaches, employing terms such as "we calculate with confidence that . . . ", without providing quantitative measures. The problems of linguistic assessment without quantification are well known (Wallsten et al. 1993). The TAR uncertainty guidance document (Moss and Schneider 2000) changed this by encouraging scientists to assign a quantitative scale to linguistic claims. The Moss and Schneider scheme was partially adopted in the TAR, and provided a forceful starting point for effective communication of uncertainty.

There are however reasons for the IPCC to extend uncertainty communication beyond the level of confidence scale. Reasonable argument can be made objecting to the use of a purely subjective scale. For one, uncertainty surrounding a few key variables in the historical record (e.g. global mean temperature change) can be assessed in a largely objective manner. In such cases, it probably makes practical sense to preserve the distinction between objective and subjective assessment. Second, some members of the scientific community might, for a variety of reasons, not be comfortable using an exclusively Bayesian approach (Giles 2002). Since the IPCC is ostensibly a consensual scientific organization it takes intellectual pluralism seriously, and accommodates differing but valid perspectives on uncertainty. Hence, using levels of confidence along with likelihood provides a useful way of combining different levels of knowledge, while satisfying the needs of a consensus process. We turn next to the question of how likelihood and confidence can best be used in combination.

2.3 Alternative 3: use levels of confidence and likelihood

In this section we discuss two schemes for combining measures of both likelihood and confidence. The first scheme uses both measures together to condition likelihoods by level of confidence. The second scheme adjusts the metric used to express likelihood according to the level of confidence.

2.3.1 Alternative 3a: simultaneous use of likelihood and levels of confidence

The simplest possible approach is to simultaneously combine likelihood and levels of confidence in communicating uncertainty. For instance, the AR4 uncertainty guidance document allows provision for likelihood to be used to communicate the probability/variability in a particular outcome (Table 2), and for the level of confidence to communicate the level of agreement associated with that likelihood (Table 3). The logic of this approach appears to be reasonable—provide likelihoods based on available data, but also communicate a subjective uncertainty about the likelihood. Upon closer inspection it becomes clear that simultaneous use of likelihood and level of confidence can cause confusion and make the already difficult challenge of communicating uncertainty even more difficult. Below we provide a description of why this may be the case.

All likelihood outcomes (of high, medium or low likelihood) with a low subjective confidence cannot be interpreted in a quantitative manner. Part of the problem with the approach is that likelihood and confidence cannot be fully separated. Likelihoods contain implicit confidence levels. When an event is said to be extremely likely (or extremely unlikely) it is implicit that we have high confidence. It wouldn't make any sense to declare that an event was extremely likely and then turn around and say that we had low confidence in that statement. For example, if we declare that it is extremely likely to rain tomorrow, but then say that we have very low confidence in that statement, that would lead to a state of confusion. People would rightly ask us how we could give such a high (near certain) likelihood to an event about which we profess to have little understanding. If we say there is a 99% chance of rain, that implies that we are nearly certain it is going to rain, which means that we must have high confidence, never low.

As we show in Table 4, interpreting uncertainty when there are two levels of imprecision is in some cases rather difficult. The table shows likelihoods conditioned by levels of confidence. First consider a statement whose likelihood is high, and subjective confidence on this likelihood is high. This entry in the table is easy to interpret and would have high likelihood. In fact, all entries in the row related to high confidence are easy to interpret. The likelihood gives an estimate of the probability of some event occurring, and the high confidence estimate tells us that the subjective error bars for that estimate are small. Thus, the likelihoods conditioned by confidence level are equivalent to the likelihoods when the level of confidence is high.

Now consider entries in the row related to low confidence, and we run into difficulties of interpretation. If low confidence translates into large error bars about the likelihood estimate, then the actual likelihood (could it be known) may bear little relationship to the estimated likelihood. Further, we encounter real problems when combining very high or very low likelihood estimates (the 'virtually certain' and 'exceptionally unlikely' from Table 2) with low confidence assessments. As pointed out earlier, very high or very low likelihoods are only meaningful when confidence is high. By allowing low confidence assignments to such estimates, confusion may be created. The simultaneous likelihood/confidence scheme allows the analyst to create contradictory combinations of likelihood and confidence.

Combinations of likelihood and confidence with medium levels of confidence are intermediate between the high and low confidence cases, which makes them somewhat ambiguous. If medium confidence implies moderate error bars, then the

Table 4 Likelihoods conditioned by levels of confidence

		likelihood low	medium	high
	low	?	?	?
confidence	medium	low–med	low–high	med–high
	high	low	medium	high

This table shows the simultaneous use of likelihoods and levels of confidence. Confidence and likelihood levels are each classified into 'low', 'medium' and 'high'. Entry in each element of the table represents how the associated confidence level modifies the likelihood. For example, when the likelihood of an outcome is low and subjective confidence levels about the science surrounding that outcome are high, then the likelihood is low. However, when likelihood of an outcome is high and the subjective confidence levels associated with that outcome are low it is impossible to interpret the likelihood in a meaningful way. Such combinations that are difficult to interpret are represented by a '?'.

likelihood estimates may be reasonable, but perhaps one category too high or low. In Table 2 then, the low likelihood estimate might correspond to a confidence conditioned likelihood in the range from low to medium likelihood. And the high likelihood estimates could correspond to conditioned likelihoods from medium to high for example. That is, medium confidence implies a spreading of the likelihood ranges, assuming that the error bars are moderate. However, if medium confidence implies larger error bars, then the same confusion that applies to the low confidence cases would apply to some degree to the medium confidence cases.

A scheme that combines estimates of confidence and likelihood is increasingly difficult to interpret the lower the estimate of confidence. In the extremes at low confidence and very high or low likelihood, the combinations make little sense. These features of the scheme would create a conundrum for analysts that may lead them to avoid low and medium confidence combinations. This could result in a bias toward expressing results with higher confidence, since it is meaningful with this scheme to present only statements associated with higher confidence. Thus an uncertainty scheme that simultaneously uses likelihood and confidence is ripe to either contradict itself or bias towards suppression of low confidence. Hence, we argue that simultaneous use of likelihood and levels of confidence can be dangerous. Below we propose a method that merges some of the positive aspects of both approaches (alternatives 1 and 2).

2.3.2 Alternative 3b: adjust likelihood scale according to warranted precision

Another approach to using both likelihoods and levels of confidence is to use a progressive scheme that articulates the basis for the assessment of each attribute (Risbey et al. 2002; Kandlikar et al. 2005). This scheme allows analysts to use a sequential process that does not treat all uncertain variables as statistically quantifiable, and provides a mechanism for communicating uncertainty at a level appropriate to existing scientific understanding. The sequential process is outlined in Table 5 and described below. The process begins by asking the analyst if a probability distribution for the outcome or variable under consideration can be provided (i.e., full likelihood information). This serves to capture either those variables for which historical data exists, or those for which there is sufficient consensus. If a pdf can be given then one moves on to the next variable of interest. For many/most variables this is not the case however, so it is necessary to have more coarse means of representing the uncertainty as well. In these cases the analyst moves down to the next level in the

Table 5 Characterizations of likelihood for a graduated range of precision levels ranging from fully specified probability distributions through to qualitative declarations of knowledge and ignorance

	Measure of likelihood	Justification
1	Full probability density function	Robust, well defended distribution
2	Bounds	Well defended percentile bounds
3	First order estimates	Order of magnitude assessment
4	Expected sign or trend	Well defended trend expectation
5	Ambiguous sign or trend	Equally plausible contrary trend expectations
6	Effective ignorance	Lacking or weakly plausible expectations

scheme. At each level down the degree of quantification (precision) is reduced. The idea is to express quantities at a level of precision commensurate with the degree of confidence with which the quantity is known. The steps in the scheme are as follows:

Step 0: Definition Define the variable or outcome to be examined and the context in which it is being examined. Though seemingly trivial, this first step is crucial in ensuring that the outcome in question has a commonly shared understanding and can be meaningfully quantified. This step also facilitates comparison of uncertainties across studies and through time.

Step 1: Full probability density function (Robust, well defended probability distribution): Is it reasonable to specify a full probability distribution for the outcome? If yes, specify the distribution. Justify your choice of distribution and 5th and 95th percentiles. Are there any processes or assumptions that would cause the 5/95 percentiles to be much wider than you have stated? This is the full likelihood description. If you cannot provide justifications for why you consider the distribution shape and 5th and 95th percentiles to be fairly robust, then move to a lower precision category (step 2).

Step 2: Bounds (Well defended bounds): Is it reasonable to specify bounds for the distribution of the outcome? If yes, specify 5th and 95th percentiles. Can you describe any processes or assumptions that could lead to broader/narrower bounds? If so, describe and revise. The choice of 5th/95th percentiles is by convention. Other ranges (e.g. 10th/90th) could also be used by different research communities as long as the choice is made clear. If the bounds are robust to assumptions, then specify your 5/95 bounds and your reasoning for placing them where you did. If you cannot provide bounds confidently then go to step 3.

Step 3: First order estimates (Order of magnitude assessment): If appropriate, specify and justify your choice of a first order estimate for the value of the variable, indicating the main assumptions behind the value given. In specifying a value, do not report more precision than is justified. For example, if the value is only known to a factor of two or an order of magnitude, then report it in those terms. In some cases, powers of ten may be appropriate; in other cases more nuanced scales may be used so long as they are declared and supported. How robust is your estimate to underlying assumptions? If it is not particularly robust to the set of assumptions or outcomes you listed, then go to step 4.

Step 4. Expected sign or trend (Well defended trend expectation): While it may not be possible to place reliable bounds or a magnitude on the expected change in a variable, you may still know something about the likely trend. Can you provide a reasonable estimate of the sign or trend (increase, decrease, no change) of the expected change? If so, give the expected trend and explain the reasoning underlying that expectation and why changes of the opposite sign or trend would generally not be expected. Describe also any conditions that could lead to a change in trend contrary to expectations. It is reasonable to include in this category changes which have a fair degree of expectation, but which are not certain. The distinction between this category and the following one is that the arguments for the expected change should be significantly more compelling or likely than those for a contrary change. If the arguments tend towards a more equal footing, then step 5 (ambiguous sign) is more appropriate.

Step 5: Ambiguous sign or trend (Equally plausible contrary trend expectations): In many cases it will not be possible to outline a definitive trend expectation. There may be plausible arguments for a change of sign or trend in either direction. If that is the case, state the opposing trends and outline the arguments on both sides. Note key uncertainties and assumptions in your arguments and how they may tip the balance in favour of one trend direction or the other. If information about the variable does not support this kind of supposition, then go to step 6.

Step 6: Effective ignorance (Lacking or weakly plausible expectations): In most cases we know quite a bit about the outcome variable. Yet despite this, we may not know much about the factors that would govern a change in the variable of the type under consideration. As such, it may be difficult to outline plausible arguments for how the variable would respond. If the arguments used to support the change in the variable are so weak as to stretch plausibility, then this category is appropriate. Selecting this category does not mean that we know nothing about the variable. Rather, it means that our knowledge of the factors governing changes in the variable in the context of interest is so weak that we are effectively ignorant in this particular regard. If this category is selected, describe any expectations, such as they are, and note problems with them.

These six steps provide a mechanism for making explicit the reasons for low/high levels of confidence based on assessments of data quality and scientific knowledge at each step. Responses can be given in progressively relaxed quantitative forms, ranging from full likelihood form (when justified) through to more qualitative characterizations as appropriate. The analyst moves down through the steps and stops when the level of confidence in the variable matches the precision available in the category. Though the method is subjective, it is transparent in that it asks the analyst to provide justifications for the form of quantification selected. Thus the reasoning is clear and explicit for others to scrutinize. The approach provides a simple, yet consistent way to use likelihood information in conjunction with subjective knowledge.

In this scheme, the form (scale) in which likelihoods are expressed is conditioned by subjective judgement (confidence) such that likelihood and confidence remain consistent with one another at low confidence levels—likelihoods are expressed in coarse quantitative form when confidence is low. This contrasts with the simultaneous likelihood/confidence scheme where likelihood is conditioned by level of confidence, but the form of likelihood expression does not change.

The determination of appropriate precision levels to express outcomes proceeds through a process of argumentation that progressively excludes over-precise and under-precise levels. In practice that determination won't always be obvious, and the analyst may wish to employ a variety of measures to assess the quality and precision of the outcome variable. One approach which appears to be well suited is to use the NUSAP scheme (Funtowicz and Ravetz 1990), which employs methods to determine the 'pedigree' and quality of the relevant data and methods used (e.g. van der Sluijs et al. 2005a,b). Any method which helps to determine the appropriate precision of outputs could be used in this regard.

3 Likelihood example

A simple example may be helpful in illustrating use of the likelihood schemes. Suppose we are asked what the likelihood is of the thermohaline circulation shutting down in response to increased greenhouse forcing of the climate (Mastrandrea and Schneider 2004). First, we need to remove any ambiguity in the question, so we would need to specify a time range over which this event might occur. In the limit to infinity, the question would be almost trivial as the circulation has shutdown in the past and would likely do so at some point in the future, greenhouse warming or not. Thus, we might limit the period to some point in time such as 2100 or until some particular CO_2 concentration is reached, say 550 ppm. Further, we need to define whether the flow simply slows down or reverses completely. Suppose we rephrase the question then as to what is the likelihood that the thermohaline circulation reduces by at least half for a stabilized CO_2 concentration of 550 ppm? Of course one would also need to specify or factor in the time path of greenhouse and other emissions and the climate sensitivity to make the question more precise and account for further uncertainties. Since we only consider hypothetical responses to the question by way of example, the full specification of uncertainties is not critical here.

In the case that the likelihood was very high that the circulation would be reduced by at least half, the answer could be phrased easily enough with either scheme. In the combined likelihood/confidence scheme (scheme 3a) one would use the likelihood scale of Table 2 to yield the answer 'virtually certain'. In order to make such a certain determination, confidence would also have to be very high. In the precision-based scheme (scheme 3b), if the underlying knowledge about thermohaline circulation responses to CO_2 were very robust, one could specify a probability distribution for the value of the flow at 550 ppm. From the distribution it would be clear in this case that the vast bulk of the probability mass favoured flows less than half the present value.

Suppose now that confidence in the assessment of thermohaline circulation changes was not high, but low. Following the earlier quote of Manning and Petit (2004), could one meaningfully specify that the outcome was virtually certain with low confidence? We suspect that this would be interpreted by many to mean that the likelihood of the likelihood (virtually certain) was low; i.e. that it was not likely to be in this category. But that is presumably not exactly what is intended, for one would do much better in that case to simply specify a more likely likelihood category. Further, the specification of likelihoods upon likelihoods conjures up the notion of an infinite regress (Funtowicz and Ravetz 1990). Presumably, what is intended here is that one has low confidence in specifying the likelihood and really means to say that the likelihood is unknown. There is no point in specifying a very precise category ($> 99\%$ chance) and then saying, "well, we don't know much about that". This would be true of every likelihood category in this case and it could be misleading to single out a single category and make the statement. Rather, one wishes to convey the uncertainty about likelihood directly and without ambiguity.

In the precision-based scheme, if confidence is low there is no point trying to provide a pdf or even percentile bounds on the value of the flow at 550 ppm. One moves down through the scheme to find an expression for the expected changes in the flow that matches the level of understanding about that. If a 'first-order estimate' could reasonably be given for the value of the flow at 550 ppm, that would be specified, together with declarations of assumptions underlying the value given. If confidence were too low to warrant that, one might simply specify the 'expected sign' for changes in the flow. There are good reasons to expect a decrease in the flow (Manabe and Stouffer 1993), and this much at least could be conveyed, along with the reasoning. This is not a likelihood in that it does not answer the question in terms of a probability of the specified outcome. But since that probability would be vacuous, it doesn't make sense to give one. One provides as much quantifiable information as the level of confidence will support. If confidence were even lower again, the alternative scheme (3b) allows for more speculative declarations of knowledge and ignorance. In the fixed-precision likelihood scheme of Table 2, one cannot retreat from specifying a probability, no matter how tenuous the knowledge base. The fixed-precision likelihood scheme thus becomes a straight-jacket for the analyst when uncertainties increase.

4 Conclusions

The sequence of quinquennial IPCC reports from 1990 to the present time provides an interesting study of the evolution of formal uncertainty communication in the climatological community. In the period prior to the first IPCC reports, subjective information about uncertainty tended to be included in ad hoc ways. Some of the first reports on climate change, such as National Research Council (1979) provided discussion of the reasonings that led them to select particular ranges (for climate sensitivity for example), though ambiguity remained about just what the ranges were supposed to represent (van der Sluijs 1997). Subsequent climate change assessments used a range of different styles to communicate qualitative and quantitative dimensions of uncertainty (National Research Council 1982, 1992; Jaeger 1988).

The earliest IPCC reports provided qualitative statements of confidence in expressing results. This process was formalized by Moss and Schneider (2000) for the third IPCC assessment. They introduced a formal scale for assessing levels of confidence, which was adapted by IPCC (2006) to express confidence in terms of the odds of being correct (Table 3). With preparation for the fourth IPCC assessment, formal expressions of confidence have now been supplemented with formal expressions of likelihood (IPCC 2006) (Table 2). This provides analysts with the ability to describe both the chance of some outcome occurring and the confidence they have in their prediction of that chance. This offers greater flexibility over uncertainty expression schemes based on only one of these dimensions. With *confidence only* schemes there is too little attention to the extraction of appropriate likelihood information for risk assessments (Dessai and Hulme 2004) and inevitable ambiguity when likelihood information is given. Schemes based on *likelihood only* fall short because they assume that all relevant uncertainty information can be communicated in likelihood terms. In practice there are large subjectivities underlying many climate assessments that are best addressed through some form of confidence assessment.

Recognizing the advantages of using both likelihood and confidence information, the IPCC AR4 has provided schemes for both these concepts in communicating uncertainties (Allen et al. 2004; IPCC 2006; Manning 2006). Some analysts will undoubtedly use these two schemes in combination. However, some combinations of likelihood and confidence (as these concepts are defined by the IPCC) are difficult to interpret. The source of the problem is that likelihood levels contain implicit confidence levels. For example, very high/low likelihoods only seem meaningful if interpreted as statements of high confidence. Yet a simple combination of the IPCC AR4 likelihood and confidence schemes theoretically allows analysts to create confusing combinations of low confidence and high likelihood. Such combinations will usually be avoided. However, by avoiding the confusing combinations in this scheme the analyst may bias uncertainty communication by under-reporting low confidence cases. This also has the effect of reducing the likelihood/confidence scheme back to a *likelihood only* scheme since only higher confidence cases are retained.

A version of the alternative scheme for reporting uncertainty given here has also been incorporated into the AR4 uncertainty guidance. The scheme described here starts with the recognition that different quantities dealt with in IPCC reports are known with differing levels of precision. One-size-fits-all precision schemes are bound to be left wanting in such circumstances. The expression of likelihood needs to be flexible enough to take into account a range of precision levels from fully quantitative pdf's to virtual ignorance of quantitative changes. Rather than having a single quantitative format with a single precision for expressing likelihood, the level of precision is relaxed as the underpinning knowledge of the quantity degrades. Full quantitative information is provided for likelihoods when it is reasonable to do so, and only then. Confidence enters into the alternative scheme via choice of precision level and by way of explanation; not formally through a labelling scheme like Table 3. That is, one has to defend the choice of level of precision, explaining reasoning, outlining assumptions, and evaluating the robustness of the choice. The levels of confidence are implicit in the choices of quantitative category and in the articulation of the factors underlying those choices. The assessment of confidence thus conditions the form in which likelihood is expressed, rather than the value itself as in conventional schemes.

The IPCC's uncertainty assessments are marked by increasing formalism, thus reducing linguistic sources of ambiguity. These developments in uncertainty communication codified in the IPCC provide a richer platform to communicate climate science for policy, though potential for confusion remains. The new formalisms are beginning to incorporate deeper forms of uncertainty, opening the door for more pluralistic conceptions of uncertainty in future assessments.

Acknowledgements This work was funded by the Australian Research Council and the Wealth from Oceans Flagship program. Helpful comments were provided by Suraje Dessai, Chris Forest, and Jeroen van der Sluijs.

References

Allen M, Booth B, Frame D, Gregory J, Kettleborough J, Smith L, Stainforth D, Stott P (2004) Observational constraints on future climate: distinguishing robust from model-dependent statements of uncertainty in climate forecasting. In: IPCC workshop on describing scientific uncertainties in climate change to support analysis of risk and of options, Ireland, 11–13 May 2004

Dessai S, Hulme M (2004) Does climate adaptation policy need probabilities? Clim Policy 4(2):107–128

Funtowicz S, Ravetz J (1990) Uncertainty and quality in science for policy. Kluwer, Dordrecht, p 229

Giles J (2002) When doubt is a sure thing. Nature 418:476–478

Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) (2001) Climate change 2001: the scientific basis. Cambridge University Press, Cambridge, p 881

Houghton JT et al (eds) (1996) Climate change 1995: the science of climate change. Cambridge University Press, Cambridge, p 572

Houghton JT, Jenkins G, Ephraums J (eds) (1990) Climate change: the IPCC scientific assessment. Cambridge University Press, Cambridge, p 365

IPCC (2006) Guidance notes for lead authors of the IPCC fourth assessment report on addressing uncertainties. Appendix in Manning 2006. Adv Clim Change Res 2:13–21

Jaeger J (1988) Developing policies for responding to climate change. Technical report, World Meteorological Organization. WCIP-1, WMO/TD-No. 225, p 53

Kandlikar M, Risbey J, Dessai S (2005) Representing and communicating deep uncertainty in climate change assessments. Comptes Rendus Geosci 337(4):443–455

Manabe S, Stouffer R (1993) Century-scale effects of increased atmospheric CO_2 on the ocean-atmosphere system. Nature 364(6434):215–218

Manning M (2006) The treatment of uncertainties in the Fourth IPCC Assessment Report. Adv Clim Change Res 2(1):13–21

Manning M, Petit M (2004) A concept paper for the AR4 cross cutting theme: uncertainties and risk. Position paper for the IPCC Risk and Uncertainty workshop, Maynooth, Ireland, May 2004

Mastrandrea M, Schneider SH (2004) Probabilistic integrated assessment of ‘dangerous’ climate change. Science 304:571–575

Moss R, Schneider SH (2000) Uncertainties in the IPCC TAR: Recommendations to lead authors for more consistent assessment and reporting. In: Pachauri R, Taniguchi T, Tanaka K (eds) Guidance papers on the cross cutting issues of the third assessment report of the IPCC. Technical report. World Meteorological Organization, Geneva, pp 33–51

Murphy J, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430:768–772

National Research Council (1979) Carbon dioxide and climate: a scientific assessment. National Academy Press, Washington, DC, p 22

National Research Council (1982) Carbon dioxide and climate: A second assessment. National Academy Press, Washington, DC, p 72

National Research Council (1992) Policy implications of greenhouse warming: mitigation, adaptation, and the science base. National Academy Press, Washington, DC, p 944

Ravetz J (1971) Scientific knowledge and its social problems. Clarendon, Oxford, p 449 (reprint: Transaction, New Brunswick, 1996)

Risbey JS (2002) Comment on Soon et al. 2002: modeling climatic effects of anthropogenic carbon dioxide emissions: unknowns and uncertainties. Clim Res 22(2):185–186

Risbey JS, Lamb PJ, Miller RL, Morgan MC, Roe GH (2002) Exploring the structure of regional climate scenarios by combining synoptic and dynamic guidance and GCM output. J Clim 15(9):1036–1050

Schrader-Frechette K (1984) Science policy, ethics, and economic methodology: some problems of technology assessment and environmental impact analysis. Reidel, Dordrecht

Shackley S, Risbey J, Stone P, Wynne B (1999) Adjusting to policy expectations in climate change modeling: an interdisciplinary study of flux adjustments in coupled atmosphere-ocean general circulation models. Clim Change 43(2):413–454

van der Sluijs J (1997) Anchoring amid uncertainty. Ph.D. Thesis, University of Utrecht, Utrecht, p 260

van der Sluijs J, Craye M, Funtowicz S, Kloprogge P, Ravetz J, Risbey J (2005a) Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: the NUSAP system. Risk Anal 25(2):481–492

van der Sluijs J, Craye M, Funtowicz S, Kloprogge P, Ravetz J, Risbey J (2005b) Experiences with the NUSAP system for multidimensional uncertainty assessment. Water Sci Technol 52(6):133–144

van der Sluijs J, Eijndhoven J, Shackley S, Wynne B (1998) Anchoring devices in science for policy: the case of consensus around climate sensitivity. Soc Stud Sci 28(2):291–323

Wallsten T, Budescu D, Zwick R (1993) Comparing the calibration and coherence of numerical and verbal probability judgments. Manage Sci 39:176–190