

Improving conveyance of uncertainties in the findings of the IPCC

Rachael Jonassen · Roger Pielke Jr.

Received: 5 November 2010 / Accepted: 1 July 2011
© Springer Science+Business Media B.V. 2011

Abstract Authors of the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) received guidance on reporting understanding, certainty and/or confidence in findings using a common language, to better communicate with decision makers. However, a review of the IPCC conducted by the InterAcademy Council (2010) found that “the guidance was not consistently followed in AR4, leading to unnecessary errors . . . the guidance was often applied to statements that are so vague they cannot be falsified. In these cases the impression was often left, quite incorrectly, that a substantive finding was being presented.” Our comprehensive and quantitative analysis of findings and associated uncertainty in the AR4 supports the IAC findings and suggests opportunities for improvement in future assessments.

1 Introduction

Authors of the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) received guidance on reporting understanding, certainty and/or confidence in findings using a common language, to better communicate with decision makers (Solomon et al. 2007; Parry et al. 2007; Metz et al. 2007; IPCC 2005; Manning and Petit 2003; Moss and Schneider 2000). The AR4 Summary for Policy Makers (SPM) presented a summary of “key findings and uncertainties” but the word ‘finding’ is not used again in the document and is not elsewhere defined (Pachauri and Reisinger 2007). However, the IPCC’s formal guidance instructed author’s to assign a defined uncertainty measure to each “key policy relevant finding.” Using this criterion,

Electronic supplementary material The online version of this article (doi:10.1007/s10584-011-0185-7) contains supplementary material, which is available to authorized users.

R. Jonassen
LMI, McLean, VA 22101, USA
e-mail: rjonassen@lmi.org

R. Pielke Jr. (✉)
Center for Science and Technology Policy Research, University of Colorado/CIRES,
Boulder, CO 80309, USA
e-mail: pielke@colorado.edu

we identified 2,744 such findings in the AR4, enabling an evaluation of the treatment of uncertainty across the WGs, and tracking the provenance of findings from the SPM back to the report chapters and ultimately, to the referenced peer-reviewed literature.

The online database of these findings (annotated with search terms for source, topic, time, uncertainty and region) that we have created can be found at clicke.lmi.org. Accompanying each finding, we supply a unique identifier that allows it to be traced to its source within AR4. An advanced search feature allows users to filter on multiple search terms and combine various levels of the same search term into a single search. Counts of findings satisfying various search terms allow users to focus on specific areas efficiently and decide when a search has narrowed to findings that should be examined in detail. Users can also back navigate from a finding to identify findings sharing a common theme. We made use of these features in the following analysis with a focus on uncertainty metrics in various parts of the IPCC AR4.

This short paper describes the treatment of uncertainty by the AR4 with a focus on differences between IPCC guidance and actual practice. We find several opportunities for improved consistency in the treatment of uncertainties across Working Groups (WGs), but also, across the various levels of reporting within each Working Group. We find inconsistencies in both instances. We conclude with recommendations for a more systematic treatment of uncertainty in future assessments.

2 IPCC AR4 guidance and practice for the conveyance of uncertainty

In its guidance to authors the IPCC recommended three scales for consistent description of uncertainty across the AR4: qualitative (understanding: amount of evidence, amount of agreement), quantitative (confidence in being correct), and probabilistic (likelihood). We searched all IPCC AR4 reports and summaries for the various terms defined by the IPCC to indicate levels of confidence and uncertainty in key findings (Table 1). The presence of these terms indicated a ‘finding’ of the IPCC AR4, and each such finding was copied verbatim to create a database. We located 2,744 such findings in the AR4, enabling an analysis of the treatment of uncertainty across the IPCC, and tracking the provenance of findings from the SPM back to the report chapters and ultimately, to the referenced literature.

Because of inconsistencies in use of terminology, we broadened our criteria of a “finding” to include the variations defined by the various WGs (e.g. “more likely than not”). Since some terms (e.g. likely) are also frequently encountered in normal written language, we inspected each potential finding to confirm the use did imply a finding. We also found misspellings and frequent interjection of stray characters (e.g. Confidence, Confi dence, Vèrylikely) that required further intelligent filtering. We also accepted certain terms¹ not sanctioned by the original guidance note.² IPCC authors are warned not to use words like “likely” in a colloquial sense, particular in “elevated” language in Executive Summaries, Summaries for Policymakers, and the Synthesis

¹ The first finding presented in the IPCC AR4 SPM uses the term ‘unequivocal,’ which is not a defined term in the Guidance Document. There is no category for expressions of “ignorance,” i.e., situations where uncertainties are simply not known.

² We did not accept such uses as “Most likely” or “Uncertain” as indicators of a finding.

Table 1 Frequency of key findings in IPCC AR4 source documents and associated measures of understanding, confidence, and likelihood

Measures of likelihood	Synthesis report	Working group 1	Working group 2	Working group 3 ^a	Grand total
Unequivocal	2	2			4
Virtually Certain (>99%)	3	4	26		33
Extremely Likely (>95%) ^b		6			6
Very Likely (>90%)	34	125	161	1	321
Likely (>66%)	51	327	744	35	1157
More Likely than not (> 50%)	3	17	5 ^c		25
About as likely as not (33% to 66%) ^d		1	3		4
Unlikely (<33%)		13	28	12	53
Very Unlikely (< 10%)	7	16	6		29
Extremely Unlikely (<5%)	1	4			5
Exceptionally Unlikely (<1%)		2	1		3
Measures of Confidence	SR	WG I	WG II	WG III	Grand Total
Very High Confidence (>90%)	7	2	106		115
High Confidence (~80%)	28	7	265		300
Medium Confidence (~50%)	14	2	125	1	142
Low Confidence (~20%)			32		32
Very Low Confidence (<10%)			1 ^e		1
Measures of Understanding	SR	WG I	WG II	WG III	Grand Total
High Agreement, Much Evidence	13	3		240	256
High Agreement, Medium Evidence	5			71	76
High Agreement, Low Evidence				18	18
Medium Agreement, High Evidence	1	7		9	17
Medium Agreement, Medium Evidence				67	67
Medium Agreement, Low Evidence				28	28
Low Agreement, High Evidence	3	12			15
Low Agreement, Medium Evidence	4	13		2	19
Low Agreement, Low Evidence		10		8	18
TOTAL	176	573	1503	492	2744
Findings within Chapters		411	1251	227	1888

^a WGIII repeats likelihood and confidence findings from Chapter 19 of WGII

^b WGI uses this likelihood measure but it is not included in the IPCC Guidance Note

^c All of these are drawn from WGI.

^d WGI substitutes the term 'more likely than not' with probability>50%. See WGI-SPM, p.3

^e This finding cautions against a *certain* inference since it would result in *very low confidence*

Report. We found enough ambiguity in usage that suggests that this is an issue worthy of reemphasis for AR5.

Each of the WGs explained independently their use of measures of uncertainty, and thus, perhaps unsurprisingly, each WG treats uncertainties in independent fashion, leading to

inconsistencies across the WGs. For example, WGIII asserts that the use of the quantitative measures is inappropriate “as human choices are concerned,”³ ⁴ ⁵ and finds (with *High Agreement and Much Evidence*) that the recommended methods of reporting uncertainty in the IPCC Guidance (i.e. Confidence and Likelihood) are not appropriate for the topics discussed in WGIII.

WGI Chapter 2 introduced a measure of uncertainty not covered in the IPCC’s guidance to authors, Level of Scientific Understanding (LOSU), modified from the Third Assessment Report (TAR).⁶ WGII and most chapters of WGI did not present levels of agreement or understanding, so presumably found other means to assess degree of consensus. However because the IPCC guidance was not followed, this information remains out of sight. The 17 findings in the Summary Report (SR) and Summary Report Summary for Policy Makers (SR-SPM) reporting level of understanding all derive from WGIII.

With the different approaches across the working groups to the expression of uncertainty, interpretation of the findings can be difficult. For instance, although the Guidance Note cautioned against over-using a mid-range category (and indeed did not specify such a category) “as a ‘safe’ option that communicates little information to the reader,” 40% of WGIII findings fall in this (medium) category. WG I further expanded the three levels by hyphenating classes (e.g. medium-low). Moreover, WGIII reported only a single finding with a corresponding measure of confidence even though the IPCC guidance recommends that findings supported by much evidence and with high agreement (over half of WGIII findings) should be accompanied by a description of uncertainty (quantitative or probabilistic).

WGIII reports probabilistic measures of uncertainty for 48 findings or one-fifth of those that might qualify. Of the 12 findings using the defined term ‘unlikely’ only the one that appears in the WGIII Technical Summary (TS) has an accompanying indication of understanding (in this case it does not rise to the level recommended in the guidance). Similarly, of the 35 ‘likely’ findings, only the 4 in the WGIII-TS have a measure of understanding attached and none of these follow the guidance, and the one ‘very likely’ finding has no measure of understanding assigned. WGIII almost exclusively uses of the first (qualitative) measure of confidence but with incomplete adherence to the guidance.

WGII used both quantitative (for 32% of its findings) and probabilistic (68%) measures of uncertainty when reporting findings and WGI used probabilistic measures almost exclusively (98%). WGII occasionally also mixed two methods of representing quantitative uncertainty within a single finding, for example, they report:

*Vulnerability is **likely** to rise due to an increase in extreme events: Economic damage from extreme weather is **very likely** to increase and provide major challenges for adaptation (**high confidence**).* [11.ES][WG2]

Confusion about the implications of having high ‘confidence in being correct’ (~80% certainty) that something is likely (>66%) would be understandable.

³ The description of uses of measures of uncertainty for WGII references a URL which does not exist <http://www.ipcc.ch/activity/uncertaintyguidancenote.pdf> (accessed 1-12-11)

⁴ Chapter 2 provides a very useful discussion of the history of the subject and choice of methods for this WG. Results presented in Table 1 suggest that only Working Group III (Mitigation) found the IPCC Guidance Note useful to express explicitly their amount of evidence. WGIII defines ‘evidence’ (WGIII-TS, p.34 and Table TS-1) and it is found in the WGIII Glossary but not in other IPCC Glossaries.

⁵ WG III TS p.34, see also WGIII-SPM Endnote 1, p.23.

⁶ The method WGI used to collapse the four-item, two-dimensional measure of understanding defined in the Guidance Note to a five-level one-dimensional LOSU measure is described in Chapter 2, p. 200 and Table 2.11.

WGI reports findings using both the likelihood and confidence language and add the level ‘more likely than not.’ This new term is used in WGI in chapters 1 (once), 9 (6 times, including four times in Table 9.4) and 11 (twice). WGI (Table 11.2) mixes likelihood (e.g. very likely, likely) and confidence (e.g. medium confidence) in the same scale.

Findings reported for each WG in Table 1 include those in each chapter and its Executive Summary as well as those found in the TS and the SPM for each WG. Generally, findings in the WG-TS (112, 154, 239)⁷ and WG-SPM (50, 98, 26) reflect findings in the WG chapters. But those findings may be stated differently or combined with other findings as they move up the hierarchy of documents of AR4 (see Table 2 and the two examples in the [Supplementary Material](#)). The Synthesis Report (SR, with 111 findings) and its SPM (with 65) draw on all three WGs’ findings and filter more severely. Ideally, measures of understanding would form one basis for such filtering and their availability would better inform the process.

Uncertainties are transformed in other ways between the chapters of the full report and the summaries. For example, an important finding on methane is somewhat disconnected from its original presentation in the full text of the report, having been greatly simplified and expressed with greater certainty than is offered in the source chapters (see the [Supplementary Material](#) for documentation). In this case, the data presented in the relevant WG chapters do not support the highest-level finding. We have found other similar examples where uncertainties were presented differently in the full text and in the summaries, representing an inconsistent presentation of knowledge at different stages of the IPCC process.

If we confine our attention to those findings that refer to the future, one can ask how many IPCC findings can be expected to become verified ultimately as being accurate? For example, if we consider findings that refer to future events with likelihood in the ‘likely’ class (i.e., >66% likelihood) then if these judgments are well calibrated then it would be appropriate to conclude that as many as a third can be expected to not occur. More generally, of the 360 findings reported in the full text of WG1 across all likelihood categories and presented with associated measures of likelihood (i.e., those summarized in Table 2 below), then based on the judgments of likelihood associated with each statement we should logically expect that about 100 of these findings (~28%) will at some point be overturned.⁸

The WGs chapters reported 21 findings about which they were ‘virtually certain’ and two of those were reported in the SR accompanied by the same measure of confidence. One of those two was then reported in the SPM with the same measure of confidence. There is value in presenting these findings in a manner that allows tracking them through the assessments directly to their origins, with a clear provenance. Such clarity would provide ready answers to such questions as:

- What is the process for changing the wording of findings as they move to successive levels of the hierarchy of documents of the AR4?
- Who decides what things that we’re virtually certain about get into the Synthesis Report and Summary for Policy Makers?
- What is the knowledge foundation on which the finding is built?
- How does knowledge and expressed certainty evolve from assessment to assessment?

⁷ Counts for each WG

⁸ This calculation assumes that each finding can be treated independently. If the findings are not independent (e.g., they are cumulative) then this calculation would result in a higher estimate.

Table 2 Examples of the number of findings with an associated likelihood identified at various levels of the hierarchy of IPCC AR4 documents

	Syn Rpt SPM	Syn Rpt	WG1 SPM	WG1 TS	Chapter totals	WG1 Chp1	WG1 Chp9
Unequivocal	1	1	1	1			
Virtually Certain	1	2	1	2	1		
Extremely Likely				1	5		3
Very Likely	14	20	17	36	72		9
Likely	17	34	21	47	259	2	63
More Likely than not	1	2	6	5			4
About as likely as not					3		
Unlikely				2	11		2
Very Unlikely	2	5	3	7	6		1
Extremely Unlikely		1	1	2	3		1
Exceptionally Unlikely							1

There are insufficient data to determine whether consistency exists across the WGs for assessing confidence or understanding. Indeed, the fact that measures of understanding are not reported by WGII, and a unique method is used in WGI Chapter Two, suggests that greater consistency may be of value. The fact that WGII reports nearly three times as many statements of uncertainty as the other WGs (Table 1) is, at least, curious. Moreover, the distribution of the findings changes greatly from one WG to another (Fig. 1). WGIII reports one finding with a measure of confidence, but 48 findings with associated likelihood (all derived from WGI).

3 Summary and discussion

The IPCC recommended three scales of uncertainty: qualitative (understanding: amount of evidence, amount of agreement), quantitative (confidence in being correct), and probabilistic (likelihood). Each of the three Working Groups (WGs) independently applied measures of uncertainty, including in each case idiosyncratic deviations from IPCC guidance. Not surprisingly, each WG's treatment of uncertainty is inconsistent with the others (IAC 2010). Others have claimed that the different treatment of uncertainties by different WGs is necessary due to the different subject matter and methodological approaches between scientific disciplines (Swart et al. 2009).

We think it unlikely that all the differences can be attributed to such disciplinary idiosyncrasies given the maturity of the field of uncertainty analysis, the effort taken to coordinate methods prior to initiating the effort, and the breadth of metrics available in the guidance note. For example, the claim that quantitative or probabilistic measures of uncertainty are inappropriate for an entire WG report comprising over 13 chapters, hundreds of pages and nearly 500 findings on topics as diverse as energy and waste management may be more a decision of individuals rather than the consensus of all authors.

The modal class of likelihood for all three WGs is 'likely' (i.e., >66% likelihood). WGI emphasizes certainties of *Very Likely* or higher in the Summary for Policy Makers (46%), Synthesis Report (36%), and the overall Synthesis Report Summary for Policy Makers (38%), but only 22% of all findings reach this level in the WGI chapters. While there may

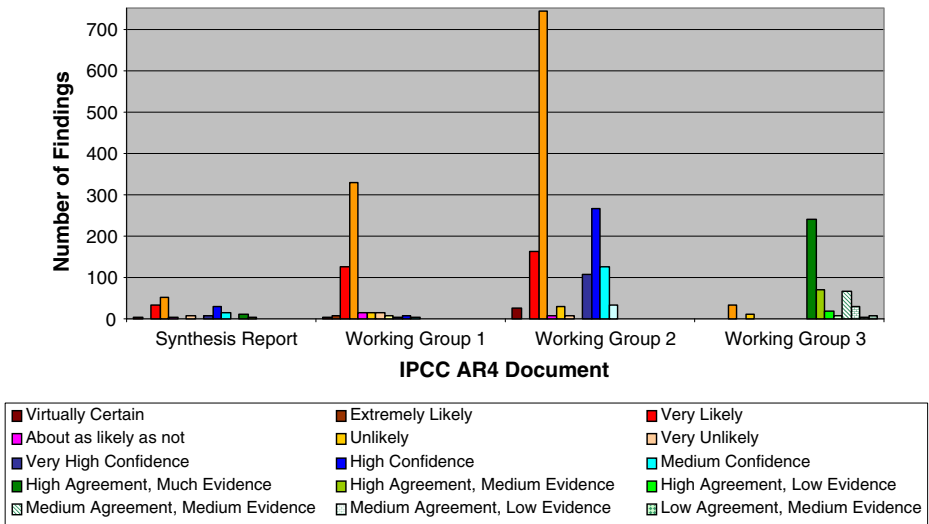


Fig. 1 Distribution of uncertainty types and levels associated with findings of the IPCC AR4 working groups and comparison to the synthesis report

be very good justifications for emphasizing greater certainties to policy makers, to the flip side of emphasizing certainties is to deemphasize uncertainties (PBL Netherlands Environment Agency 2010). It may be the case that decision makers could benefit from understanding what knowledge is more uncertain. Given that the IPCC has observed that action on climate change necessitates decision making under uncertainty, the different emphasis on certainties presented in the various summary documents as compared to the full text of the reports may actually hinder decision making.

The IAC recommended that IPCC authors “should provide a traceable account of how they arrived at their ratings for level of scientific understanding and likelihood that an outcome will occur” (IAC 2010, p. 37). To illustrate such a “traceable account” the IPCC could create a database of findings, associated uncertainty level, regional relevance, topic and time period (past present, future), along with links to source material, and information on the process of documenting and selecting findings through the hierarchy of documents. The database should include all findings from all reports and link findings across ARs where appropriate. This would show the evolution of knowledge and certainties. We have created a database of the AR4 findings to illustrate one manner in which this might be done in future assessments. This database is available online at clিকে.lmi.org.

It is not enough to simply provide guidance on expression of uncertainty, or even to systematically follow that guidance consistently. To fully incorporate the implications of uncertainties in findings, IPCC must specify how to combine uncertainty metrics at each stage of the synthesis process and move to a system of independent tracking of findings and their attendant uncertainties during the writing process. The method of representing uncertainty should be robust enough to encompass all findings. We recognize this is a learning process; we can expect changes in the method along the way. Such learning should be a planned part of the process. Certainly the large improvement in treatment of uncertainty through the sequence of IPCC Assessment Reports speaks highly of the effort made and dedication to improvement.

Although the IPCC has made enormous contributions and set an important example for global assessment of a vexing problem of immense ramifications, there remain clear

opportunities for improvement in documenting findings and specifying uncertainties. We recommend more care in the definition and determination of uncertainty, more clarity in identifying and presenting findings and a more systematic approach in the entire process, especially from assessment to assessment. We also suggest an independent, dedicated group to monitor the process, evaluate findings as they are presented and track their fate. This would include tracking the relationship of findings and attendant uncertainties that pass up the hierarchy of documents within AR5. Strict rules for expressing uncertainty in findings that are derived from (possibly multiple) other findings are needed (see, e.g., the second example in the Supplementary Material).

It is not the purpose of this note to discuss other, related scientific assessments of climate change knowledge; but, we do note that our preliminary analysis of the U.S. Global Change Research Program Synthesis and Assessment Products suggests a far less systematic application of the guidance supplied to authors of those documents and far less consistent application of the defined terms. We believe that the concerns we have expressed here, and the resulting recommendations, apply more broadly than the IPCC process.

Acknowledgements The authors thank Ms. Julia Kalloz (LMI) for supervising the collection and organization of all the IPCC AR4 findings. Ms. Kalloz also performed much of the initial analysis of finding frequency and oversaw indexing and tagging findings with keywords. Ms. Shaun Lehman (Virginia Tech), Ms. Kristen Cleven (Columbia University) and Ms. Rebecca Freeburn (George Washington University) assisted in collecting the IPCC findings while interns at LMI. Ms. Kelli Canada (LMI intern), Mr. Mark Gross (LMI), Mr. Aman Luthra (LMI), Ms. Lia Marroquin (LMI), and Mr. Daniel Jackson (LMI) assisted tagging findings with search terms. Cathy Becker (LMI) advised on frequency analysis. Mr. Greg Wilson (LMI), Mr. David Manzollilo (LMI), Mr. David Shepherd (LMI), and Mr. Reddy Enugu (LMI) developed the web-based version of CliCKE. We wish to acknowledge the support of the LMI Research Institute for multiple grants to develop the CliCKE database, the underlying research and preparation of this manuscript. Ami Nacu-Schmidt aided in the preparation of the manuscript. We thank an anonymous reviewer for useful comments that improved the text.

References

- InterAcademy Council (2010) Review of the IPCC. <http://reviewipcc.interacademycouncil.net>, p. 36
- IPCC (2005) Guidance note for lead authors of the IPCC fourth assessment report on addressing uncertainties. Intergovernmental Panel on Climate Change, Geneva, p 4
- Manning MR, Petit M (2003) A concept paper for the AR4 cross cutting theme: uncertainties and risk. In: Manning MR, Petit M, Easterling D, Murphy J, Patwardhan A, Rogner HH, Swart R, Yohe G (eds) IPCC workshop on describing scientific uncertainties in climate change to support analysis of risk and of options: workshop report. Intergovernmental Panel on Climate Change, Geneva
- Metz B, Davidson OR, Bosch PR, Dave R, Meyer LA (2007) Contribution of working group III to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom
- Moss R, Schneider S (2000) Uncertainties. In: Pachauri R, Taniguchi T, Tanaka K (eds) Guidance papers on the cross cutting issues of the third assessment report of the IPCC. Intergovernmental Panel on Climate Change, Geneva
- PBL Netherlands Environment Agency (2010) Assessing an IPCC assessment: an analysis of statements on projected regional impacts in the 2007 report. <http://www.pbl.nl/en/publications/2010/Assessing-an-IPCC-assessment.-An-analysis-of-statements-on-projected-regional-impacts-in-the-2007-report>
- Pachauri RK, Reisinger A (2007) Climate change 2007: synthesis report. Contribution of working groups I, II and III to the fourth assessment report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland, 104 pp

- Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (2007) Contribution of working group II to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom
- Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (2007) Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom
- Swart R, Bernstein L, Ha-Duong M, Petersen A (2009) Agreeing to disagree: uncertainty management in assessing climate change, impacts and responses by the IPCC. *Clim Chang* 92(1–2):1–29. doi:[10.1007/s10584-008-9444-7](https://doi.org/10.1007/s10584-008-9444-7)