Climatic Change (2025) 178:106
https://doi.org/10.1007/s10584-025-03935-2

RESPONSE

Navigating confidence–precision trade-offs in assessment

Richard Bradley[1] · **Casey Helgeson**[2,3] · **Brian Hill**[4,5]

Received: 7 April 2025 / Accepted: 13 April 2025

Abstract
In this reply, we address a comment on our paper "Combining probability with qualitative degree-of-certainty metrics in assessment" (Helgeson et al. Clim Change 149(3):517–525, 2018). Our original paper proposes an incremental systematization of confidence and likelihood language used by the Intergovernmental Panel on Climate Change (IPCC). Our goals were to improve consistency across findings and support use of confidence judgments in decision making. The comment critiques our proposal and recommends against its adoption. We argue that this recommendation is based on two misunderstandings. The first concerns trading off confidence against the precision of a finding (our proposal endorses and systematizes the practice). We defend this practice and attribute opposition to an overzealous Bayesianism inapt for the IPCC context. The second misunderstanding concerns our purported commitment to a specific procedure for producing confidence judgements. We clarify that our proposal makes no such commitment. We also note, contrary to the comment's claim, that a version of the procedure in question has been used in the IPCC's Sixth Assessment Report.

Keywords Confidence · Likelihood · Precision · Uncertainty · IPCC

This reply refers to the comment available online at https://doi.org/10.1007/s10584-025-03931-6.

✉ Richard Bradley
r.bradley@lse.ac.uk

Casey Helgeson
helgeson@psu.edu

Brian Hill
hill@hec.fr

[1] Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, London, UK

[2] Earth and Environmental Systems Institute, Penn State University, University Park, PA, USA

[3] Department of Philosophy and Rock Ethics Institute, Penn State University, University Park, PA, USA

[4] GREGHEC, CNRS, Jouy-en-Josas, France

[5] Department of Economics and Decision Sciences, HEC Paris, Jouy-en-Josas, France

Published online: 14 May 2025

The starting point of our paper "Combining probability with qualitative degree-of-certainty metrics in assessment" (Helgeson et al. 2018) is the framework for consistent treatment of uncertainties used in the fifth and sixth assessment reports (AR5 and AR6) of the Intergovernmental Panel on Climate Change (IPCC). That framework employs both interval-valued likelihood judgements and ordinal confidence judgements. On some well-studied, policy-relevant topics – and primarily in Working Group I (WGI)– the two are combined, with confidence levels qualifying statements of likelihood. Our aim was to interpret and clarify the relationship between likelihood and confidence when combined in this way. We proposed a formal model of the confidence–likelihood relationship that provides support for consistent usage in uncertainty communication, downstream modelling, and decision making.

The comment builds a complex argument against our proposal. The argument begins with the question of whether we interpret *likelihood judgements* as subjective or objective claims. Interesting though this question of interpretation may be, it is not one we attempted to address in our paper, and, as we shall explain, we are not persuaded that the answer to it makes a good deal of difference to our model of confidence. Consistent with guidance for IPCC authors (Mastrandrea et al. 2010, 2011), we are deliberately agnostic on this interpretive question. The comment finds this agnosticism untenable but for reasons that derive from two key misunderstandings. The first concerns the propriety of the confidence trade-offs sanctioned by our proposal; the second whether it rests on a specific procedure for producing confidence judgements. We respond to both of these below. First, however, we briefly reiterate the logic of our original proposal.

IPCC confidence levels primarily qualify claims or findings. Our approach begins from the observation that more precise (i.e., logically stronger) claims should not be held with greater confidence than less precise ones. For example, AR6 reports *medium confidence* that "the Atlantic Meridional Overturning Circulation will not collapse abruptly before 2100" (B.3.3, IPCC 2023). It would be incoherent, we claim, to also report only *low confidence* in no collapse before 2050 (although one could coherently report higher confidence in this event). In addition to being highly intuitive, this confidence–precision constraint also follows from the assessment of confidence in terms of underlying evidence and agreement (Mastrandrea et al. 2010): evidence for lasting past 2100 is also evidence for lasting past 2050, and weakening a claim (from >2100 to >2050) makes agreement on that claim easier. Moreover, the two comprehensive reviews of uncertainty language in IPCC reports both find evidence that authors implicitly recognise such a confidence–precision trade-off and do their best to navigate it appropriately (Janzwood 2020; Mach et al. 2017).

Where confidence qualifies statements of likelihood, the preceding logic applies to the precision of probability intervals. This opens the door to reporting likelihood for the same quantity or event at multiple confidence levels—e.g., *medium confidence* that event A is *very likely* (90–100%) and *high confidence* it is *more likely than not* (50–100%). Here confidence trades off with the precision of the likelihood assessment. As above, the idea is that the same underlying body of science (observations, studies, theories, etc.) may sometimes support the validity of cautious, less informative claims more than that of sharper, stronger claims. One way for a claim to be less informative is to assign a broader probability range—i.e., to put weaker constraints on the probability of a given outcome. Our proposed model captures such trade-offs using nested sets of probability distributions—one for each confidence level (see Bradley et al. 2017; Helgeson et al. 2018 for details).

The commentary denies that different confidence levels can be intelligibly applied to multiple likelihood statements on the same quantity or event, saying that this interpretation of confidence is "unconvincing, as the agent can increase confidence by lowering the precision of their credal judgments without any change in the underlying evidence" (footnote 8). Whilst no defence is provided for the "unconvincing" verdict, a clue may come from the earlier argument that, if the probability density functions (pdfs) grouped into our nested sets are interpreted as Bayesian posteriors, then:

> "… any subset of pdfs whatsoever is based on the available evidence, and hence the very same evidence. So if we interpret confidence as expressing something like Keynes' "weight of evidence", then any subset of pdfs would have to be assigned the same confidence level."

The worries expressed in this argument are unfounded. At the risk of repeating points from above, *even if* the evidence underlying multiple sets of pdfs is the same, the confidence warranted in the wider probability intervals corresponding to the superset can be greater than that warranted in the more precise intervals corresponding to a subset. This could be, for example, in virtue of greater agreement underpinning the former. The contrary view in the comment may reflect a confusion surrounding how likelihood claims and the confidence judgements qualifying them can be based on the same evidence and yet be different. The answer is that they reflect *different* aspects of the evidence. Likelihood judgements track the balance of evidence for or against an empirical claim; confidence reflects the type, amount, quality, and consistency of the evidence (its weight) that can be brought to bear and which underpins the balancing exercise.

Underlying this argument could be the commentary's adoption of a philosophical view – dubbed *Bayesianism* – on which statements of probability *by definition* comprehensively wrap in all available evidential considerations. On this view, there can be no space for characteristics of underlying evidence to play any *additional* role, and elevating any probability estimate over another in terms of validity or trustworthiness becomes a theoretical non-starter. But there are well known reasons why this view cannot do justice to the IPCC's two-dimensional uncertainty language (documented in Bradley et al. 2017). Indeed, distilled to its core, the central dilemma proposed by the comment in terms of objective vs subjective probabilities is a rehearsal of these arguments. No surprise, then, that problems arise when trying to fit our proposal into a strict Bayesian schema, for that schema cannot support *any* interpretation of the IPCC uncertainty language.

Our interpretation of confidence and likelihood fits into an alternative framework for uncertainty representation and decision making, set out and defended in Hill (2013, 2019a) and Bradley (2017). It is explicitly aimed at overcoming the weaknesses of the Bayesian vision of uncertainty representation and rational decision when it comes to ambiguity or deep uncertainty and is, we claim, the appropriate normative replacement for Bayesianism in such cases. The framework makes no use of the distinction between objective and subjective probabilities underlying the dilemma outlined in Section 3.1 of the comment and need not commit to a single interpretation. Indeed, since the sets of probabilities are simply a representation of the likelihood judgements – not something that is conceptually prior to them – they inherit whatever interpretation is given to those likelihood judgements.

Our second key point concerns the comment's assertion (Section 3.2) that we prescribe a *procedure* for determining confidence-qualified likelihood judgements. This is a misinterpretation. We propose no specific procedure for attributing confidence and are largely neutral on how authors should arrive at likelihood judgements. Indeed, it is an advantage of a formally grounded interpretation of the kind we propose that it supports a variety of procedures suitable for different contexts. For instance, traditional approaches to assigning probabilities include expert elicitation, aggregation of expert judgements, and (Bayesian) updating on new information; corresponding approaches are available for confidence as well (Hill 2019b, 2022, 2024). Where useful, IPCC authors could draw on these or other procedures to help implement confidence reporting consistent with our proposal.

By contrast, the comment links our proposal to one specific procedure and then critiques that procedure to undermine the proposal. The attributed procedure draws on an illustration of our model of confidence used to interpret example findings from AR5 (Bradley et al. (2017). The illustration involves a set of probability distributions which are partitioned into tiers of plausibility. Likelihood judgements consistent with *all* distributions (even less trusted ones) are qualified with *higher confidence* while likelihood judgements consistent with *only* narrower subsets of distributions are qualified with *lower confidence*. Although we did not intend the illustration as a procedure, one could imagine arriving at confidence judgments by manually assembling such a mathematical structure. For instance, the probability distributions might come from different modelling studies, ordered by the plausibility of the assumptions involved in the respective models. The comment reads a procedure like this into our proposal, disparaging it as "significantly disconnected from current IPCC practices."

While our proposal does not stand or fall with this or any specific procedure, we note that something rather similar was in fact used in the AR6. As explained in Kopp et al. (2023), authors of the chapter on Ocean, Cryosphere, and Sea Level Change (Fox-Kemper et al. 2021) examined collections of alternative pdfs for global mean sea level (GMSL) at future points in time (e.g., 2100) under alternative shared socioeconomic pathways. These pdfs came from studies that project sea-level rise using different methods, models, and assumptions. The chapter authors partitioned the pdfs into two groups based on confidence in the approaches that produced them, leading to one group "incorporating only processes in whose quantification there is at least medium confidence" and another group that "incorporated processes in whose quantification there is low confidence" (pp1295–6). Hinkel et al. (2019) do something similar, both cases resulting in "tiered imprecise probability distributions of different levels of confidence" (Kopp et al. 2023).

Interpreting the resulting findings requires some care due to differences in the *objects*, so to speak, of the confidence judgements. When discussing future sea level, the AR6 chapter applies confidence language primarily to *underlying methods* rather than directly to resulting GMSL ranges or their likelihood (pp1295–99). One example that does both is a key finding for 2050. Across all scenarios, medium-confidence methods project GMSL falling "between 0.1 and 0.4 m (5th–95th percentile range)" (p1296). Including low-confidence methods raises the upper end of this range to 0.6 m, but only under RCP8.5. "On the basis of these studies," the authors conclude, "we therefore have *high confidence* that GMSL in 2050 will be between 0.1 and 0.4 m higher than in 1995–2014 under low- and moderate-emissions scenarios, and between 0.1 and 0.6 m under high-emissions scenarios." In other words, the authors apply *high confidence* to the *broader* outcome range that encompasses the results of both low- and medium-confidence methods (consistent with the logic of our proposal).

Here and elsewhere, the low-confidence methods produce a *wider* spread of pdfs than medium-confidence methods. As a result, an inverse relationship can arise between confidence levels applied to methods or models and those applied directly to associated outcome or likelihood ranges. Where the chapter authors communicate confidence levels *in methods*, they use confidence language to signal that the broader estimate of likely outcomes is tainted by low-confidence methods. In contrast, where they assess confidence directly in the outcome or likelihood range (as above), they use confidence language to signal the *conservativeness* of the broader range in the sense that it remains accurate *even if the results of low-confidence methods turn out correct*. Both messages are valid and important and the difference is one of emphasis.

Putting aside potential confusion arising from the disparate objects of confidence judgements, the end result in terms of guidance for decision making is similar in spirit regardless of how confidence language is deployed to qualify reported *likely* ranges. The model of decision making that partially motivates our proposal tells decision makers to use the broader range of probability estimates (labelled as higher confidence on our proposal) where the *stakes* of a decision are comparatively high (Bradley 2017; Hill 2013, 2019a). Hinkel et al. (2019), Oppenheimer et al. (2019), and Kopp et al. (2023) arrive at a similar place, endorsing use of the broader range of probability estimates (however confidence labels are used) in cases of low *risk tolerance* (Oppenheimer et al., Kopp et al.) or *uncertainty tolerance* (Hinkel et al.). Clarifying how uncertainty assessments in the form of confidence-qualified likelihood judgements could serve to support risk-sensitive decision making was a key motivation for the development of our model. The comment gives no reason to think that it cannot live up to this ambition.

Acknowledgements For discussion of a draft, we thank James Doss-Gollin, Robert Kopp, Adam Pollack, and Vivek Srikrishnan.

Author contributions Equal.

Funding This work is supported by Penn State's Rock Ethics Institute and the National Science Foundation as part of the Megalopolitan Coastal Transformation Hub (MACH) under NSF award ICER-2103754 (MACH contribution number 77).

Data availability Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Consent to publish All authors consent to publication.

Competing interests None.

References

Bradley R, Helgeson C, Hill B (2017) Climate change assessments: confidence, probability, and decision. Philos Sci 84(3):500–522. https://doi.org/10.1086/692145

Bradley R (2017) Decision theory with a human face. Cambridge University Press.https://doi.org/10.1017/9780511760105

Fox-Kemper B, Hewitt HT, Xiao C, Aðalgeirsdóttir G, Drijfhout SS, Edwards TL, Golledge NR, Hemer M, Kopp RE, Krinner G, Mix A, Notz D, Nowicki S, Nurhati IS, Ruiz L, Sallée J-B, Slangen ABA, Yu AY (2021) Ocean, cryosphere and sea level change. In: Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou AB (eds) Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press, pp 1211–1362 https://doi.org/10.1017/9781009157896.011

Helgeson C, Bradley R, Hill B (2018) Combining probability with qualitative degree-of-certainty metrics in assessment. Clim Change 149(3):517–525. https://doi.org/10.1007/s10584-018-2247-6

Hill B (2013) Confidence and decision. Games Econom Behav 82:675–692. https://doi.org/10.1016/j.geb.2013.09.009

Hill B (2019a) Confidence in beliefs and rational decision making. Econ Philos. https://doi.org/10.1017/S0266267118000214

Hill B (2022) Updating confidence in beliefs. J Econ Theory 199(105209):105209. https://doi.org/10.1016/j.jet.2021.105209

Hill B (2024) Confidence, consensus and aggregation. SSRN Electron J. https://doi.org/10.2139/ssrn.4387641

Hill B (2019b) Confidence in belief, weight of evidence and uncertainty reporting. Int Symp Imprecise Probabilities Appl 103:235–245. https://proceedings.mlr.press/v103/hill19a.html

Hinkel J, Church JA, Gregory JM, Lambert E, Le Cozannet G, Lowe J, McInnes KL, Nicholls RJ, Pol TD, Wal R (2019) Meeting user needs for sea level rise information: a decision analysis perspective. Earth's Future 7(3):320–337. https://doi.org/10.1029/2018ef001071

IPCC (2023) Summary for policymakers. In: Core Writing Team, Lee H, Romero J (eds) Climate change 2023: synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change. IPCC, pp 1–34. https://doi.org/10.59327/IPCC/AR6-9789291691647.001

Janzwood S (2020) Confident, likely, or both? The implementation of the uncertainty language framework in IPCC special reports. Clim Change 162(3):1655–1675. https://doi.org/10.1007/s10584-020-02746-x

Kopp RE, Oppenheimer M, O'Reilly JL, Drijfhout SS, Edwards TL, Fox-Kemper B, Garner GG, Golledge NR, Hermans THJ, Hewitt HT, Horton BP, Krinner G, Notz D, Nowicki S, Palmer MD, Slangen ABA, Xiao C (2023) Communicating future sea-level rise uncertainty and ambiguity to assessment users. Nat Clim Chang 13(7):648–660. https://doi.org/10.1038/s41558-023-01691-8

Mach KJ, Mastrandrea MD, Freeman PT, Field CB (2017) Unleashing expert judgment in assessment. Glob Environ Change: Hum Policy Dimens 44:1–14. https://doi.org/10.1016/j.gloenvcha.2017.02.005

Mastrandrea MD, Mach KJ, Plattner G-K, Edenhofer O, Stocker TF, Field CB, Ebi KL, Matschoss PR (2011) The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups. Clim Change 108(4):675–691. https://doi.org/10.1007/s10584-011-0178-6

Mastrandrea MD, Field CB, Stocker TF, Edenhofer O, Ebi KL, Frame DJ, Held H, Kriegler E, Mach KJ, Matschoss PR, Plattner G-K, Yohe GW, Zwiers AFW (2010) Guidance note for lLead aAuthors of the IPCC fFifth aAssessment rReport on cConsistent tTreatment of uUncertainties. Intergovernmental Panel on Climate Change (IPCC). Available at https://www.ipcc.ch

Oppenheimer M, Glavovic BC, Hinkel J, van de Wal R, Magnan AK, Abd-Elgawad A, Cai R, Cifuentes-Jara M, DeConto RM, Ghosh T, Hay J, Isla F, Marzeion B, Meyssignac B, Sebesvari AZ (2019) Sea level rise and implications for low-lying islands, coasts and communities. In: Pörtner H-O, Roberts DC, Masson-Delmotte V, Zhai P, Tignor M, Poloczanska E, Mintenbeck K, Alegría A, Nicolai M, Okem A, Petzold J, Rama B, Weyer NM (eds) IPCC special report on the ocean and cryosphere in a changing climate. Cambridge University Press, pp 321–445. https://doi.org/10.1017/9781009157964.006