

Burning Embers: Towards more transparent and robust climate change risk assessments

Authors and Affiliations

Zinta Zommers - United Nations Office for Disaster Risk Reduction, Platz der Vereinte Nationen 53113 Bonn, Germany;

Philippe Marbaix - UCLouvain, Georges Lemaître Centre for Earth and Climate Research, Louvain-la-Neuve, Belgium;

Andreas Fischlin - ETH Zurich, Terrestrial Systems Ecology, Institute of Biogeochemistry and Pollutant Dynamics, Department of Environmental Systems Sciences, Universitaetstrasse 16, 8092 Zürich, Switzerland;

Zelina Z. Ibrahim - Universiti Putra Malaysia, Faculty of Forestry and Environment, Serdang, Malaysia;

Sean Grant - Indiana University, Richard M. Fairbanks School of Public Health, Indianapolis, IN, USA, 46204;

Alexandre K. Magnan - Institute for Sustainable Development and International Relations (Sciences Po, 27 rue Saint Guillaume, 75007 Paris, France) & UMR LIENSs 7266 (La Rochelle University-CNRS, 2 rue Olympe de Gouges, 17000 La Rochelle, France);

Hans-Otto Pörtner- Alfred Wegener Institute, Section Integrative Ecophysiology, Am Handelshafen 12, 27570 Bremerhaven, Germany;

Mark Howden - Climate Change Institute, Australia National University, Canberra, Australia;

Katherine Calvin - Pacific Northwest National Laboratory, College Park, MD 20740 USA;

Koko Warner - United Nations Climate Change Secretariat, Subprogramme on Climate Impacts, Vulnerabilities and Risks, Platz der Vereinte Nationen 53113 Bonn, Germany;

Wim Thiery - Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Brussels, Belgium;

Zita Sebesvari - United Nations University, Institute for Environment and Human Security, Bonn, Germany;

Edouard L. Davin - ETH Zurich, Institute for Atmospheric and Climate Science, Switzerland;

Jason P. Evans - University of New South Wales, Climate Change Research Centre and ARC centre of excellence for Climate Extremes, Sydney, NSW, Australia

Cynthia Rosenzweig - NASA Goddard Institute for Space Studies, New York, NY, USA & Center for Climate Systems Research, Columbia University, New York, NY, USA;

Brian C. O'Neill - University of Denver, Josef Korbel School of International Studies, Denver, CO, USA;

Anand Patwardhan - University of Maryland, School of Public Policy, College Park, USA;

Rachel Warren - University of East Anglia, Tyndall Centre, Norwich, Norfolk, United Kingdom;

Maarten van Aalst - University of Twente, Faculty of Geo-information science and earth observation (ITC), Enschede, the Netherlands, and Red Cross Red Crescent Climate Centre, The Hague, the Netherlands

Margot Hulbert - University of Regina, Centre for the Study of Science and Innovation Policy, Johnson-Shoyama Graduate School of Public Policy, Regina, Sk, Canada

Corresponding Author: zommers@un.org

Abstract

The Intergovernmental Panel on Climate Change (IPCC) reports provide policy-relevant insights about climate impacts, vulnerabilities, and adaptation through a process of peer-reviewed literature assessments underpinned by expert judgement. An iconic output from these assessments is the “burning embers” diagram. Burning embers were first used in the Third Assessment Report to visualize reasons for concern, which aggregate climate change-related impacts and risks to various systems and sectors. In these diagrams, colour transitions show changes in the assessed level of risk to humans and ecosystems as a function of climate change

indicated by increasing levels of global mean temperature. Here we review the history and evolution of the burning embers diagram and the associated reason for concern framework describing the methodological approach to risk assessment and construction of the diagram. We highlight relevant expert elicitation methods in other fields and detail recent methodological advances in IPCC Special Reports aiming to reduce the risk of potential bias in the expert judgement, improve the transparency of the assessment, and better account for adaptation. We recommend the use of an explicit, transparent, and systematic process of expert elicitation in the production of burning embers diagrams to enhance their quality, comparability, and credibility in future assessments.

Key Points

- Since 2001, the IPCC has used the reasons for concern framework and burning embers diagrams to assess and communicate risks from increasing global mean temperature on human and natural systems.
- The framework and diagram are developed using expert judgement, based on available information about climate impacts as arising from climate hazards, exposure and vulnerability including the capacity to adapt.
- While assessment methods and diagram design have been broadly retained across reports, risk levels at given temperatures have generally increased with each assessment based on new, more comprehensive science.
- Structured expert elicitation methods that specify the process for selecting experts, providing external information, eliciting individual and consensus judgments, facilitating group interaction can reduce bias and increase result reproducibility.
- Recent IPCC Special Reports introduced formal protocols and standardized metrics to elicit risk thresholds. Despite challenges such as limited time or funds, these changes contributed to transparency and reliability.
- Further development and use of standardized and transparent methods to elicit risk thresholds and build burning ember diagrams could continue to increase the robustness and credibility of these assessments.

1. Introduction

Climate change impacts are unfolding at a pace, intensity, and scale that captures increasing public attention¹. A key policy objective to limit human interference in Earth's climate system is expressed in Article 2 of the United Nations Framework Convention on Climate Change (UNFCCC): to stabilize greenhouse gas concentrations at a level that will prevent dangerous anthropogenic interference with the climate system². However, it has been a challenge to identify precisely when anthropogenic interference becomes "dangerous" because of uncertainties in model projections, sectoral and regional differences in impacts, and because understanding of what is "dangerous" depends on value judgements³⁻⁶.

To help facilitate judgements about risks, the Intergovernmental Panel on Climate Change (IPCC) introduced the "reasons for concern" (RFC) framework and the associated colour-coded figures, commonly called "burning ember diagrams", in the Working Group II (WGII) contribution to the Third Assessment Report (TAR)^{7,8}. In the RFC framework, risks from global mean temperature rise are aggregated into five categories: 1) risks to unique and threatened

systems; 2) risks associated with extreme weather events; 3) risks associated with the distribution of impacts; 4) risks associated with global aggregate impacts; and 5) risks associated with large-scale singular events⁹. Risk levels are identified using expert judgement, based on available information about relevant hazards, exposure, vulnerability, capacity to adapt, and impacts¹⁰. RFCs are visualized as “burning embers”, where different risk levels are expressed with different colours and uncertainties about precise changes conveyed through graded colour transitions (Figure 1).

The RFC framework and burning embers diagram are used as tools for communicating risks related to anthropogenic climate change in a way that can support both public discussions and policy decisions. For example, the Paris Agreement goal to hold global average temperature “well below 2 °C”¹¹, was informed by IPCC reports which indicated increasingly high risks and limited adaptive capacity beyond warming of 1.5 or 2°C^{5,12}. Accurate understanding of climate-related risks continues to be critical for decision making, as governments update their Nationally Determined Contributions and work to develop national and sub-national disaster risk reduction strategies mandated by the Sendai Framework for Disaster Risk Reduction. Indeed, at UNFCCC COP 25, governments called for increased support on risk assessments, including long-term risks from climate change impacts (2/CMA.2 para.26 (a)). The sixth IPCC assessment report (AR6) will include a chapter on risks across sectors and regions which is expected to update the RFCs and possibly the burning embers diagrams¹³. The IPCC 6th Assessment cycle including its Special Reports is expected to inform upcoming climate policy processes, such as the first Global Stocktake of the Paris Agreement in 2023.

The RFC framework and burning embers diagrams have also been the subject of criticism^{14–19}. It has been argued that their production should be more systematic, transparent, and comparable across reports²⁰. Early assessments of RFCs or burning embers figures did not detail methods, particularly with regard to the assignment of risk thresholds. These gaps led to questions regarding the reliability of the methods and reproducibility of the burning embers, potentially undermining confidence in the robustness of the RFC framework and its use for policy making. Some critics have argued that the use of expert judgement in IPCC reports has been “application dependent, controversial and implemented differently across chapters and reports”¹⁵. Critics call for a more formalized approach integrating numerical modelling with expert judgement and indicating the full range of assessments made as well as any disagreements²¹. Much guidance material has been produced by IPCC and the research community to strengthen IPCC assessments on aspects such as the role of expert judgement and uncertainty language^{22–24}. Nevertheless, questions about the RFCs and the burning embers remain.

In this review, we trace the evolution of the RFC framework and burning embers diagrams, focusing specifically on changes in methods and expert elicitation processes. A full analysis of expert elicitation methods is beyond the scope of this paper and is available elsewhere^{25–32}. Here, we highlight relevant lessons from other fields and document efforts to develop a more systematic approach to estimating the risk transitions that occurred in the 2019 IPCC Special Report on Climate Change and Land (SRCCCL)³³ and 2019 Special Report on the Oceans and Cryosphere in a Changing Climate (SROCC)³⁴. Methods are presented and discussed together with possible improvements to the formal process of developing expert consensus in future IPCC cycles. The review is based on published literature and on the experience of authors of this

paper who have participated in IPCC assessments from TAR onwards – including the assessment of RFC's and construction of burning embers diagrams.

2. History

To begin, we outline the history of the RFC framework and burning embers diagram, highlighting changes in methods, design, and risk thresholds.

2.1 Emergence of framework and figure

RFCs and the burning embers figure first appeared in Chapter 19, *Vulnerability to Climate Change and Reasons for Concern: A Synthesis*, of the WGII contribution to TAR⁷. The multidisciplinary author team, which included physical, biophysical, and social scientists, reviewed observed impacts from climate change detailed in the available peer-reviewed literature. They debated the level of risk that climate change presented in areas relevant to the interpretation of dangerous anthropogenic interference. This led to the insight that a single metric or measure would not be sufficient to capture the diversity of risks relevant to policy discussions. Besides global economic gains or losses, reasons for concern about climate change included the numbers of people affected by projected climate hazards as well as effects on ecosystems³⁵⁻³⁷. While detailed development of the IPCC risk framing did not occur until AR5, reasons for concern were considered as relevant attributes of the 'risks' of climate change.

Early drafts of the chapter included three lines of evidence, later renamed Reasons for Concern, in relation to increases in global mean temperature (GMT): impacts on unique and threatened systems (such as tropical glaciers, coral reefs, or indigenous communities); global aggregated impacts (such as net damages for market and non-market sectors at the global scale); and the probability of large-scale singular events (such as the shutdown of the North Atlantic thermohaline circulation or the collapse of the West Antarctic ice sheet). Two other RFCs in relation to increases in global mean temperature were subsequently added: distribution of impacts (including the heightened vulnerability of developing countries); and probability of extreme climate events (including floods, soil moisture deficits, tropical storms) ^{7,38-40}.

The Chapter 19 TAR author team mapped future risks along a GMT scale and used the consensus of their expert judgments to determine transitions between different levels of risk as a function of levels of GMT. GMT increase was assessed in increments of whole degrees as small ($\sim 2^{\circ}\text{C}$), medium ($2-4^{\circ}\text{C}$), and large ($>3^{\circ}\text{C}$)⁷. The evaluation process was conducted among the authors during lead author meetings, and in response to the rounds of expert and government review. Judgments were based on the authors' expert assessment of the literature, but this process was not conducted nor documented systematically.

The GMT rise above pre-industrial levels was decided to be the most appropriate metric against which risk levels could be assessed, due to the availability of impact literature which could be related to the GMT, and the fact that the GMT could also be related to GHG concentrations¹⁷. GMT as used in TAR, and in this paper, refers to global mean surface temperature change as used also by all WGs in the AR5. Inevitably, use of GMT as a metric involved assessing and synthesising literature in which similar levels of warming were reached at different future dates. This meant that it was not possible to directly detail issues such as rates of change and

the dynamics of potential adaptation, in particular the issue that slower rates of warming could allow for a greater level of adaptation and hence lower risks.

While a grey-scale version of the burning embers was included in Chapter 19, the first colour rendition was published in 2001 WGII TAR Summary for Policy Makers (Fig. 1)⁴¹, using colours similar to a traffic light arrangement. Traffic light colours have been used to inform decision making in various contexts: in fisheries management⁴²⁻⁴⁴, the food industry^{45,46}, in health⁴⁷, geosciences^{48,49}, and in organisational performance⁵⁰. Already in 1990, in discussions on possible long-term targets for climate change at the Second World Climate Conference (1990), traffic light colours were used to represent different levels of increased risk due to climate change⁵¹⁻⁵³. In the TAR, authors decided against the use of the green colour to avoid the impression of safety or absence of risk^{51,54}. Instead, white was used to indicate neutral or small negative or positive impacts. Yellow was used to indicate negative impacts for some systems or low risks that could become evident at the denoted increase in GMT. While red was used to indicate negative impacts or risks that could be more widespread or greater in magnitude⁵⁵. Gradients of shading across the colours reflected the fact that “change in global mean temperature should be taken as an approximate indicator of when impacts might occur; they are not intended to define absolute thresholds or to describe all relevant aspects of climate-change impacts”^{7,56}. Uncertainties about transitions resulted from various sources including global warming projections, changes in adaptation or social vulnerability over time, nature of the impact assessments, and expert judgement itself.

2.2. Evolution of Methods and Design

Since the TAR, the RFC framework and burning embers diagrams have been used in nearly all IPCC Working Group II Reports as well as in literature outside of IPCC reports⁵⁷⁻⁶⁰ (Figure 2). Methods for constructing burning embers, as well as their design, have been elaborated over time.

Chapter 19 of the IPCC’s Fourth Assessment Report (AR4) included a section updating and revising the RFC assessment⁶¹. In particular, chapter 19 identified a number of criteria that could be used to justify the choice of specific reasons for concern⁶². Criteria included: (i) magnitude of impacts; (ii) timing of impacts; (iii) persistence and reversibility of impacts; (iv) potential for adaptation; (v) distributional aspects of impacts and vulnerabilities; (vi) likelihood (estimates of uncertainty) of impacts and vulnerabilities and confidence in those estimates; and (vii) importance of the system(s) at risk. Only the first six criteria were applied to the assessment, as criteria seven was deemed to be in the realm of the judgements of policy-makers³.

As in TAR, results were based on expert judgement about findings in recent literature and IPCC reports. The assessment relied on the judgements of a different set of authors. Only two of the nineteen AR4 Chapter 19 authors had also contributed to Chapter 19 of TAR. For each RFC the evidence of impacts was compiled, and a conclusion made on the location of the transition between different levels of risk as a function of GMT. For example, judgements about “Risks of extreme weather events” were based on the literature presented in the AR4 WGI report⁶³ on increasing frequency and intensity of weather extremes and on the observed impacts from such extreme weather events presented in other chapters in the AR4 WGII report^{64,65}. The RFC

framework and a burning embers representation was the subject of considerable debate during the report preparation, with some voices arguing that both the choice of the RFCs and the assignment of risk levels to different levels of warming involved too high a degree of expert judgment¹⁴. Results were ultimately considered robust and comparable to TAR because “judgments were vetted by three rounds of IPCC review and approved in the SPM of both the AR4 WGII and Synthesis Reports”⁶².

Although early versions of the AR4 report included a placeholder for a burning embers diagram, no updated diagram was included in either Chapter 19 or in the Summary for Policymakers of WGII. An updated burning embers figure was later published in a peer-reviewed academic journal⁶² while variations of burning embers figures appeared in different chapters of WGII AR4 main report^{66,67}. For the first time, risks to specific sectors and risks under different warming trajectories were explored. A figure in Chapter 4 (Figure 4.4)⁶⁵, and a simplified variant in the Technical Summary (Figure TS.6)⁶⁸, depicts risks to ecosystems against different levels of global mean annual temperature rise according to two different global warming trajectories. Colour transitions corresponding to those of burning ember diagrams were applied to visualise risks for categories of impacts and ecosystems in general⁶⁶. A figure in Chapter 11 depicts vulnerability of key sectors to climate change against different emissions scenarios, highlighting possible coping and adaptive capacity at different levels of temperature increase⁶⁷.

RFCs were published in the Fifth Assessment report (AR5) WGII Chapter 19⁶⁹ and the SPM⁷⁰. AR5 built on a new risk framework adapted from the 2012 IPCC Special Report on “Managing the Risk of Extreme Events and Disasters to Advance Climate Change Adaptation” (SREX)^{71,72}. This framework more formally expresses risks as determined by the combination of climate stressors or “hazards”, ecosystems’ and societies’ exposure to these hazards, and the vulnerability of these ecosystems and societies. Thus, a given frequency and magnitude of extreme climate events, for example, does not by itself indicate a particular level of risk. Future conditions of exposure and vulnerability, as well as the ability to adapt, can vary widely, and substantially affect the risk associated with the hazards. In AR5 and previous RFC assessments, exposure and vulnerability were incorporated implicitly in risk judgments to the extent that the underlying literature took it into account. Autonomous adaptation as well as limits to adaptation (related to the purple zone of the burning embers, see below) were also considered^{9,68}. The literature was not extensive enough at that time to support explicit differentiation of risks along adaptation pathways or specific dimensions of vulnerability and exposure. However, AR5 authors suggested that future risk judgments in RFCs and burning ember diagrams should be a function not only of global mean surface temperature change, but also of levels of exposure, vulnerability and adaptation^{9,69}.

An initial assessment of risk transitions for each RFC was made by Chapter 19 authors and based primarily on literature. These authors then confirmed or revised their judgments based on input from other chapters in the WGII report. That input took the form of a novel identification of “key risks,” defined as climate-related risks with the potential to become severe and where there is limited ability to adapt. WGII authors identified 102 key risks in their domains, which Chapter 19 authors then grouped into eight overarching categories of key risks and mapped them to the five RFCs. The assessments of these groups of key risks were then considered in the Chapter 19 judgments of the location of risk transitions for each RFC. As before, risk transitions were identified and assessed via group consensus. However the more

systematized approach of using key risks to inform RFCs was deemed to allow consistency and transparency as well as integration of different types of evidence⁹.

Further, in AR5 and a subsequent academic paper, confidence levels were introduced for transitions based on the type, extent, and agreement of evidence⁹. These confidence levels complemented the visual representation of uncertainty as blurred boundaries between risk transitions in the burning embers diagrams. The confidence levels helped differentiate higher from lower confidence judgments on the transitions between risk levels. For example, transitions from undetected to moderate risk for RFCs 1 and 2 (unique and threatened system, and extreme weather events) were made with high confidence based in part on a growing literature on the detection and attribution of impacts. In contrast, judgments about risk transitions for RFC 5 (large-scale singular events) were made with medium confidence, given substantial uncertainty in the projection of the timing of ice-sheet loss⁹.

While the design and aesthetics of the burning embers has been broadly retained across successive IPCC reports, variations have occurred. In concert with the growing literature on detection and attribution of impacts due to climate change, the colour scale of the embers was slightly modified in AR5⁶⁹. Thus, the white colour now indicated undetectable risk with no impacts “detectable and attributable to climate change”. The yellow colour indicated moderate risk where “impacts are both detectable and attributable to climate change with at least medium confidence”. The red colour indicated high risk, where risks of severe and widespread impacts are judged to be high on one or more of the specific criteria for key risks. A new colour, purple, was introduced to indicate very high risk, where all specific criteria for key risks were at very high levels, including irreversibility of an impact and exceedance of adaptation limits.

Recent Special Reports introduced refinements and variations in the burning ember diagram and RFC framework. The 2018 IPCC Special Report on Global Warming of 1.5 Degrees (SR1.5)⁷³ contains sector-specific embers along with aggregated RFC embers. Given the remit of SR1.5, risks were only assessed up to relatively low levels of global warming, focusing on present warming of +1.0°C and future warming such as +1.5°C and +2°C above pre-industrial levels. The two additional Special Reports in the AR6 cycle integrated an analysis of the role of changes in vulnerability, exposure, and adaptation for climate risk more explicitly. Building on previous recommendations and illustrations^{9,69}, the SRCCL considered differences in vulnerability (including adaptive capacity) and exposure to risk using the emerging literature based on the Shared Socioeconomic Pathways^{74–77}. SROCC introduced a burning embers diagram illustrating risk reduction in the presence of specific effective adaptation-related response⁷⁸. Other metrics instead of GMT, such as sea-level rise or atmospheric CO₂, were also used on the vertical axis of burning embers figure representing innovations that first appeared in the AR5 Synthesis report (Figure 2).

2.3. Climate Risk Thresholds

Alongside the evolution of the RFC framework and burning embers diagram has come a corresponding shift in the quantification of climate risk transitions. In most cases, the risk level at a given temperature has increased with each subsequent assessment, especially between the TAR and AR4⁷⁹(Figure 3). A comparison study of the RFCs from TAR with those based on AR4 ,

which summarized the change in the risk assessment with a single, indicative number, found that the mean of the demarcation centres between moderate and high risks decreased from 2.85°C in TAR to 1.5°C in AR4⁶⁰. A comprehensive determination of the causes of these changes in judgments would require analysis beyond the scope of this paper, but major features of the changes suggest that advances in science (including detection and attribution) and broadening of the available literature explains most of the differences between assessments⁶⁹. Smith et al (2009) indicates that, for the RFC3 (distribution of impacts) new knowledge synthesized in AR4 provided a better identification of systems, sectors, and regions that are particularly at risk, especially in Africa. Similarly, more information on changes in extremes and their impacts led to the increased risk reported in RFC2 (extreme events) in AR4.⁶²

Across RFC categories, a striking change occurred in judgments of the temperature at which risks associated with large-scale singular events (such as ice-sheet collapse) become high. This has fallen from about 5.5 °C (above pre-industrial) in the TAR to below 2°C in SR1.5 following new findings in climate science⁸⁰⁻⁸⁵. For example, AR5 Chapter 19 indicates that the main cause of change since AR4 was new evidence about ice-sheet loss during the last interglacial period, at no more than 2°C average global warming above pre-industrial⁶⁹. Adaptation to associated sea-level rise was deemed to be possible if ice-sheet loss occurs slowly, such as over a millennium⁹. However, since AR5, new observations suggest that the West Antarctic Ice sheet is already in the early stages of Marine Ice-Sheet instability⁸⁶. SR1.5 also considered findings about the slowdown of the Atlantic Meridional Overturning Current, the El Niño-Southern Oscillation (ENSO) and the role of the Southern Ocean in the global carbon cycle, concluding that risk levels at lower temperatures had increased⁸⁶.

In contrast, a number of risk transitions have remained relatively stable across multiple reports, such as the transition to medium and to high risk for the RFC on unique and threatened systems and for the RFC on distribution of impacts. New literature is still important in such cases when it provides more confidence in judgments through broader physical, ecological, and socioeconomic evidence compared to previous assessments. For example, between AR4 and AR5 literature provided new insight on how ocean acidification and warming together increase long-term coral degradation⁸⁷.

In some cases, the level of risk at a given temperature has decreased slightly in subsequent reports. For example, risk levels associated with extreme events appear at somewhat higher global warming levels in AR5 compared to AR4. At least two factors may have contributed to such changes: refinement of the framework, with clearer criteria for judging risk, and more precision in the consideration of temperature levels associated with risks. In particular, the TAR mainly refers to risk estimates for broad temperature ranges (observed past, < 2°C above 1990, 2°C-3°C, > 3°C), and it further clarifies that temperature should be taken as approximate indications of impacts, not as absolute thresholds⁷. Thus the information from the TAR was both more limited and expressed with less detail, indicating that small differences with the following reports should not be over-interpreted. Moreover, the TAR associated white areas with “no or virtually neutral impact or risk,”⁵⁶ while the AR5 refined the definition for the transition between white and yellow by adding the specific criterion that there is at “least medium confidence that impacts associated with a given risk are both detectable and attributable to climate change”⁶⁹. The requirement of attribution of impacts to climate change in AR5 has

probably also contributed to the judgment of less climate change-related risk from extreme events at low levels of warming.

These results highlight the potential use of the RFC framework and burning embers diagram to evaluate trends in risks over time. It is also clear that, to allow for comparisons, consistent and fully reproducible techniques must be used in rendering the ember diagrams, including choice of colours. This review of RFC and burning embers history indicates that assessed risk levels mainly evolved in response to new research findings, but also that the conceptual framework was supplemented and adapted over time. Continued use of the RFC framework and burning embers diagram requires enhanced and sustained attention to how the increasing amount of knowledge is taken into account to assess risks. More clarity on the details of the evolution of the conceptual framework and more rigor in the assessment methodology are needed, while considering that it would be preferable for results to be comparable to earlier ones.

3. Expert elicitation: lessons from other fields

Given the critical role of expert judgement in the RFC framework and construction of burning embers, it is prudent to explore how expert elicitation is conducted in other well-established disciplines. For example, expert elicitation is used frequently in health sciences^{25,27,32} when insufficient empirical evidence exists to inform clinical recommendations, parameters of decision analytic models, research priorities, quality indicators, or best practices in research⁸⁸⁻⁹⁶. Well-known approaches for structured, formal expert elicitation include the Classical Method⁹⁷, Consensus Development Conference⁹⁸, the Delphi⁹⁹ method, the Nominal Group Technique¹⁰⁰ and the SHeffield ELicitation Framework¹⁰¹ (Box 1). All of these approaches share several common practices for recruiting experts, preparing the elicitation exercise, eliciting and aggregating expert judgments, transforming individual judgments into data useful for analyses and aggregation, and providing feedback to experts and the wider scientific and policy communities^{26-28,32,102}. However, these formal expert elicitation methods vary in the specific design features that have been developed to facilitate the performance of experts in providing accurate, reliable, and replicable judgments (see Table 1).

Reviews of expert elicitation approaches have not identified any one standard methodology as inherently superior to others, but rather discuss which approaches may be more or less appropriate for a given context in their traditional form^{27,102,103}. Aspects of elicitation to consider for a given exercise include the data collection technique (e.g., Likert ratings, parameter values), elicitation mode (e.g., in-person, online, hybrid), and process for synthesizing individual elicitations into a group judgment (e.g., statistical aggregation, group facilitator discretion).

BOX 1: Overview of Formal Expert Elicitation Methods

The *Classical Model* of structured elicitation scores experts on their performance against empirical data for known parameters, using their performance to create and validate combinations of all expert judgments on the unknown variables of interest^{104,105}. Calibration questions are used to assess the statistical accuracy and information provided by each expert. Performance-based weights are then used for combining the expert judgments on the unknown variables of interest.

The *Consensus Development Conference* involves an open meeting over several days of a selected group of experts⁹⁸. This provides a public forum for the discussion of issues on the topic of interest. Stakeholders external to the expert group make presentations that are considered by the expert group until they reach consensus on a decision. Both the public and private sessions of the consensus development conference are chaired ³².

The *Delphi method* provides a structured, systematic communication approach for experts to independently and anonymously provide their initial judgments about the topic of inquiry. There are iterative rounds of feedback and modification of judgments based on the views of other experts on the panel⁹⁶. The final group "consensus" ²⁵ is from statistical aggregation of the individual responses.

The *Nominal Group Technique* aims to structure interaction within a group or committee with differing views¹⁰⁰. Experts first record their ideas independently and privately. After collating these ideas, the facilitator then lists one idea from each expert in front of the group in a "round-robin" fashion until all ideas have been listed and discussed. Each expert privately records their judgments for each idea until discussion ceases. Lastly, the individual expert judgements are aggregated statistically to derive the group judgement. This technique allows more ideas to be expressed and elaborated due to the initial brainstorming and following discussion of all generated ideas³².

The *Sheffield Elicitation Framework (SHELF)* elicits probability distributions for uncertain quantities from a group of experts to inform policy decisions ¹⁰¹. SHELF typically involves a face-to-face meeting with a small group (approximately 6-10) of experts led by a trained facilitator. All members are aware of each other's responses, and the group discusses modifications to a probability distribution aggregated from their individual responses until consensus is reached. Guidance and templates for pre-elicitation, elicitation, facilitation, and achieving group consensus judgments based on the perspective of a "rational impartial observer" are publicly available.

Table 1. Comparison of different expert elicitation approaches ^{26,102,105,106}

Requirement	Classical Model	Consensus Development Conference	Delphi Method	Nominal Group Technique	SHELF
Pre-elicitation training on probabilities and uncertainty	Yes	No	No	No	Yes
Primary mode of expert elicitation	Mail or phone questionnaire	In-person group discussion	Mail or online questionnaire	In-person questionnaire	Online or in-person questionnaire
Individual judgments elicited privately	Yes	No	Yes	Yes	Yes
Controlled feedback provided to participants	No	No	Yes	Yes	Yes

Multiple rounds of individual elicitation	No	No	Yes	Yes	No
Direct contact among experts	No	Yes	No	Yes	Yes
Statistical aggregation of individual elicitation into group judgment	Yes	No	Yes	Yes	Yes

Notes: Comparisons are based on the traditional or default approach to implementing each approach. Specific groups can and do make modifications in actual implementation.

Expert elicitation processes can take considerable effort on the part of the researcher and requires the commitment of the respondents. As a result, there are several considerations that need to be taken into account. While face-to-face discussions provide an opportunity for experts to examine disagreements in-depth¹⁰⁷ and take ownership of the material²⁷, they also are more costly in time and money. In addition, facilitators of expert elicitation exercises need to consider potential biases that can result from the type of respondents selected, the type of preparatory material provided, the elicitation questions and method of analysis, and specific research design. For example, anchoring subsequent questions to answers given to the first question, accessing the easiest to retrieve memory to answer questions, and lowering probabilities through range-frequency compromise are only a few of the psychological biases found in expert elicitation¹⁰⁶. Careful consideration and reporting of sources of bias is therefore required^{97,106,108}.

Despite these challenges, structured expert elicitation approaches are increasingly used in a variety of fields to fill knowledge gaps or where insufficient research or data is available¹⁰⁹⁻¹¹². Authors have argued for the application of the expert elicitation technique, as well as the increased use of systematic reviews, in environmental and climate sciences¹¹³⁻¹¹⁶. Procedures for a formal expert elicitation have been outlined, combining individual and collective deliberations¹¹⁷. Expert judgement techniques have already been used in relation to global climate change to estimate climate sensitivity¹¹⁸, tipping points in the climate system¹¹⁹, and future sea-level rise¹¹².

4. Advances in Expert Elicitation Methods in Recent IPCC Reports

Responding to earlier critiques of the standardization and rigor of the RFC framework and burning embers diagrams, Special Reports in the IPCC AR6 cycle^{33,34,78,120} have incorporated elements of expert elicitation protocols. The SRCLL focused on documenting and standardizing the expert elicitation process, while SROCC developed a standardized scoring system for risk threshold judgements.

4.1. Methods in the Special Report on Climate Change and Land (SRCLL)

The SRCLL sought to identify risks to humans and ecosystems from climate change interactions with land processes. To address these issues, a wide range of literature related to climate

change, land-use change, and socio-economic development pathways was assessed. A systematic approach was needed to characterize the risks reflected in this wider literature and to account for sources of uncertainty and variability, including professional biases inherent in individual expert's interpretation of literature.

Figure 4 outlines the broad steps used in the SRCCL. A full description is provided in this paper's supplementary material as well as in supplementary information to SRCCL Chapter 7¹²⁰. An expert elicitation process based on design features commonly employed in the Delphi and SHELF methods was used to combine the benefits of both individual and collective deliberations. To make the methodology more transparent, a protocol for eliciting expert opinions was developed including an *a priori* plan for analysing the data (i.e., a pre-analysis plan), an explicit sampling frame and eligibility criteria for selecting experts to participate in the process. Eight SRCCL authors participated in the elicitation process, representing different chapters of the SRCCL report, different regions, genders, and disciplinary backgrounds. This diversity helped decrease the impact of individual subjective biases on risk transitions and confidence levels¹²¹.

In the first step, over 300 journal articles referenced in SRCCL chapters and beyond were reviewed to extract quantitative and qualitative information about past and future impacts of climate change, socioeconomic pathways, and land use scenarios on humans and ecosystems. Evidence from each of these journal articles was added to a shared database. Using this database and considering criteria used in other parts of the report¹²², experts agreed on thresholds for each risk level. For example, a risk was considered moderate if under 1 million people or between 50 and 300 million hectares were likely to be adversely affected.

The expert elicitation took place in three rounds. In the first round, for every ember, experts provided a quantitative judgement of the GMT levels (upper bound, lower bound, and best estimate of location of transition) corresponding to each of the three risk level transitions along with reasoning for these judgements. These results were anonymised, transitions were aggregated, plotted to show the spread of results (Figure 5), and then shared with the full group of experts. In the second round, experts had the opportunity to revise their quantitative assessment and rationale. This typically involved re-examining the literature for transitions where the initial judgment differed from other experts. Results were again compiled, anonymised and shared with experts. The third round consisted of a group discussion with a facilitator, who ensured that due consideration was given to differing evidence-based viewpoints among panelists¹²³. In the group conversation, an expert would present an ember, describing their choice of transition and citing the literature that supports it. Each transition was discussed until consensus was reached. The anonymised results and facilitated discussion diminished the risk of any one expert dominating the consensus process¹²⁴.

A few factors during the group discussion helped experts refine the transition ranges and converge towards consensus. First, "very high risk" requires that "the ability to adapt is limited"⁶⁹. Agreeing on what constituted adaptation and what constituted a limit to adaptation narrowed the uncertainty range for the transition to very high risk. Second, the assessment involved a large number of studies reporting a wide variety of indicators and methods (e.g. detection and attribution, biophysical models, economic models, etc.). The group discussion led to a better understanding of how to account for diverse sources in judgements. Lastly,

uncertainty in the risk transitions is represented through both the width of the temperature transition and the confidence level of the transition. Narrow transitions may be more informative for policy-making. However, narrowing a transition range typically comes at the cost of reducing the confidence level associated with the transition. A compromise had to be found during the discussion to minimize the width of the transition while maintaining the highest possible confidence level.

4.2. Methods in the Special Report on Oceans and Cryosphere (SROCC)

While the SRCCL used a structured process to improve the robustness and traceability of the risk assessments, the SROCC report used standardized metrics to evaluate end-of-century risks from sea-level rise. The assessment also considered risks under specific adaptation pathways. Chapter 4⁷⁸ of SROCC considered four illustrative geographies covering a wide range of low-lying coastal situations across different latitudes, hemispheres, development contexts, and urban/rural settings: resource-rich coastal cities, large tropical agricultural deltas, urban atoll islands, and Arctic communities. Nine metrics were used as proxies for the components of the IPCC risk framework: i.e. hazard (coastal flooding; coastal erosion; and salinization of groundwater lenses, soils and surface waters), and exposure and vulnerability of ecosystems and people (density of assets; and degree of degradation of natural buffer ecosystems). As an innovation, four generic types of responses to sea-level rise (hard engineered coastal defences; restoration of degraded ecosystems; relocation of people and assets; and limiting subsidence) were incorporated into the assessment, making it possible to explore risk levels and transitions by 2100 under low-to-moderate and maximum adaptation.

Figure 6 outlines the broad steps used in the SROCC by Chapter 4. A full description is provided in Supplementary Information to SROCC Chapter 4 and further summarized in Supplementary Information here. The assessment method relied on a scoring system and expert judgment by eight chapter and external contributing authors, but without implementing a formal elicitation method as done in the SRCCL. The scoring system was designed to assess the relative contribution of each of the nine metrics to risk at present and by 2100 against three sea-level rise scenarios: mean RCP2.6 (+43cm), mean RCP8.5 (+84cm), and the upper *likely* range of RCP8.5 (+110cm). The scoring exercise relied on the existing literature and considered well documented specific real-world case studies, e.g. New York City (USA), Rotterdam (The Netherlands) and Shanghai (China) for the resource-rich coastal cities; and Male' (Maldives), South Tarawa (Kiribati) and Fongafale (Tuvalu) for the urban atoll islands.

For each illustrative region and each metric, the scores were aggregated by sea-level rise scenario to highlight two risk levels in 2100, one under a “low-to-moderate response” scenario (i.e., with small additional efforts in adaptation compared to today) and one under a “maximum potential response” scenario. “Maximum potential response” in this context referred to an ambitious and effective combination of both incremental and transformational adaptation (e.g., population relocation), assuming minimal financial, social, and political barriers. Such distinction was made by associating positive scores to the hazard and exposure-vulnerability metrics, as they increase risk, and negative scores to the adaptation response metrics, as they decrease risk.

The full range of theoretical aggregated scores (min.-max., from 0 to 75) represented the full

range of the IPCC risk colouring language (from white to deep purple) to highlight nine scoring levels: Undetectable, Undetectable to Moderate, Moderate, Moderate to High, High, High to Very High, Very High, Very High to Extremely High, Extremely High. With this, a specific risk level has been assigned to each aggregated score, producing the final burning embers and risk transitions.

4.3 Strengths and limitations of SRCCL and SROCC approaches

The use of formal expert elicitation processes and the development of scoring systems for judgements may have increased the transparency and robustness of the RFC framework and burning embers figures in a variety of ways. Methodological techniques used in SRCCL such as a common database of literature, anonymous judgements with justification, and use of multiple expert elicitation rounds, may have reduced potential biases such as pressure to conform to dominant individuals or anchoring in individual opinions^{29,123,125}. The scoring process in SROCC, and the articulation of metrics and judgement criteria and publication of outcome in a detailed supplement increased transparency and standardization.

Despite this progress, a variety of limitations remain. Both SRCCL and SROCC assessments could further benefit from the involvement of a greater diversity of experts. Studies have shown that heterogeneity, or a diversity of profiles and areas of expertise, in expert groups may lead to better performance compared to a more homogenous group¹²¹. Common social and psychological biases in human decision-making, such as groupthink¹²⁶ and overconfidence¹²⁷ may limit performance in homogenous groups. The SRCCL could have more clearly articulated metrics with which to evaluate risk thresholds. In turn, SROCC could have used multiple rounds of independent evaluations to better record spread and changes in judgements. Additionally, the relationship between risk and selected SROCC metrics and aggregated scores will need to be further validated.

Ultimately, the selection of an optimal expert elicitation technique within the IPCC process will need to balance the confidential nature of the process, finite funds, limited time for meetings, coordination challenges related to geographic distribution of authors across different time zones¹²⁸, the broad scope of relevant literature, the integration of different forms of evidence and different values¹²⁹. Future IPCC assessments may wish to build on, or combine, the advances made in the SRCCL and SROCC. For example, a set of experts could use a scoring system, as in SROCC, to estimate the contribution of different human-related drivers to risk levels under various climate change scenarios. Then, as in SRCCL, multiple independent assessment rounds could be used in combination with collective deliberations to discuss individual experts' underlying rationales and get consensus on final risk scores. This approach could be particularly useful for evaluating risks from climate change to human wellbeing or security – areas in which identifying temperature related thresholds for risk could be difficult given both the state of the literature and context-specificities. Future assessments could alternatively draw from other relevant expert elicitation approaches. Either way, we believe that adding structured design features to expert elicitation should be encouraged, and the advantages and disadvantages of different approaches further investigated.

5. Summary and future recommendations

The RFC framework and burning embers figure are key components of the IPCC's risk assessment process. The RFC framework helps aggregate climate-related risks into easily understood and policy-relevant categories, while the burning embers figure communicates risks using a common colour system and scale. This framework and iconic image have played a role in public policy and discourse⁶. While methods for analysis and design elements have been broadly retained across successive IPCC reports, changes have been made over time including the consideration of key risks, altered use of colours and the addition of confidence judgements. This review indicates that the risk level at a given temperature has generally increased with each subsequent assessment. It is critical to ensure these changes are driven by new science and not by methodological variations between reports or author bias. The SRCCL and SROCC added innovations to the expert elicitation process to strengthen scientific robustness and credibility.

For further enhanced usefulness in IPCC assessments, authors of IPCC reports may wish to identify and apply a standardized expert elicitation protocol to unify anonymous judgements with group discussions. Whatever method is selected, at a minimum, the protocol should specify the process for providing external information, eliciting individual judgments, facilitating group interaction, and developing consensus judgments. Assessment objectivity can be ensured by having an independent moderator, neither a chapter author nor a participant in expert elicitation, facilitate discussions. Assessment transparency can be strengthened by making workflows open, including tables with data from relevant literature, scoring and by pre-publishing the protocol⁷⁵. Authors may also wish to provide clear criteria or narratives for risk transitions used in ember diagrams. Maintaining consistent risk thresholds and metrics across chapters and reports would help strengthen communication about changes in risk levels for the decision-makers as well as the general public. Additionally, clear design protocols should be created for burning embers figure creation, including the use of standardized colours for risk levels and a standardized format to indicate confidence levels, as well as specific requirements on translating numerical risk estimates into ember graphs, preferably using a standard computer code or program.

Future risk assessments should include consideration of regional risks, the impacts of socio-economic pathways on risk and diversified adaptation scenarios and dynamics. Demand for regionally specific RFC or embers has been growing¹³⁰. The SRCCL and the SROCC laid some foundations for such analysis by including illustrative geographies or embers for specific latitudes. The SRCCL paved the way for the consideration of the influence of socioeconomic pathways on risk, while SROCC considered the relationship between different adaptation measures and risk. Assessing the potential benefits of adaptation in terms of risk reduction is key to progressively informing important and emerging policy concerns such as limits to adaptation, residual risks and benefits to be expected from both mitigation and adaptation. The influence of the rate of climate change on risks and adaptation potential could also benefit from further investigation. Further methodological improvement is needed to evaluate human adaptation, adaptation capacity relative to the degree of climate change, and adaptation limits including evolutionary and socio-cultural factors.

Going forward, the expert judgment-based risk assessments should explicitly document how they consider changes in ecological and anthropogenic exposure and vulnerability (driven by socioeconomic development as well as adaptation or mitigation responses), in addition to changes in physical climate drivers of risk related to rising global mean temperatures (as

provided by IPCC WGI). One important contribution to this integration is provided by new impact modelling initiatives that can streamline the development of burning embers across the RFCs and contribute to tracking their evolution in time. The Inter-Sectoral Impact Model Intercomparison Project (ISIMIP)¹³¹ for example, aims at harmonising impact projections across a range of sectors by providing consistent socioeconomic and meteorological forcing data and a unified simulation protocol for multi-model assessments. The ongoing phase (ISIMIP2b)¹³² provides a unique opportunity for supporting quantitative risk assessments for the AR6 cycle. Acknowledging that no one modelling approach is free of shortcomings and that models are unable to capture all dynamic processes related to impacts and adaptation, it will be important to continue to consider other sources of evidence of changing risks and effectiveness of various options for risk management.

Finally, there have been calls to better communicate the likelihood of future impacts to help policy makers set priorities on actions^{20,133}. Some attempts have been made to do this in the past. For example, Figure SPM.10 in AR5 links the levels of risk by GMT with over time cumulative anthropogenic CO₂ emissions, illustrating emissions reductions required to keep GMT, and risks, below certain levels¹³⁴. However, finding ways to clearly visualize the severity of impacts, the likelihood, and rate of change or temperature remains a challenge. Working Group I and Working Group II may wish to collaborate and identify ways to further communicate the probability of potential impacts, considering that risk associated with low likelihood but severe impacts cannot be ignored.

The run-up to the first global stocktake of the Paris Agreement in 2023 and the next round of UNFCCC's "periodic review of the long-term global goal under the Convention and of overall progress towards achieving it" provide an opportune moment for on-going work in all these areas. Increasing the reproducibility and credibility of risk assessments, and improving clarity and communication of the results, can facilitate more effective global, regional and local decision-making about consequences of dangerous anthropogenic interference with the climate system. Having more credible and traceable information about risk levels associated with different levels and rates of warming, and different socioeconomic pathways or adaptation measures, policymakers will be better equipped to make decisions about mitigation, adaptation and disaster risk reduction. The IPCC has made significant progress in this area, with innovations to the reasons for concern framework and burning embers figures in recent Special Reports. We hope this review will enable future assessments to continue to strengthen expert elicitation, thus allowing information to aid a range of climate policy decisions to minimize, and preferably avert, dangerous anthropogenic climate change.

Figure Captions:

Figure 1. The first representation of burning embers and the Reasons for Concern framework from the IPCC Third Assessment Report Technical Summary⁵⁶. The figure depicts impacts or risks from climate change by reason for concern. Each row corresponds to a reason for concern, and shades correspond to severity of impact or risk. White means no or virtually neutral impact or risk, yellow means somewhat negative impacts or low risks, and red means more negative impacts or higher risks. An identical figure, in grey scale, was included in Chapter 19 of the Working Group II Contribution to the Third Assessment Report. Iterations of

burning embers in subsequent IPCC reports are vertical with a column representing each reason for concern.

Figure 2. The evolution of the burning embers diagram in IPCC reports and other publications. Images under the left panel were published in IPCC assessments while images in the right panel were published in other scientific literature. The year for publication is indicated in the middle. The images are illustrative examples of the application of the embers concept and in most cases omit parts of the original figure. Variations in design and content have occurred since the Third Assessment Report, including the introduction of embers for specific sectors, regions in the Fourth Assessment Report and the use of axis aside from global mean temperature starting in the Fifth Assessment Report. In the Sixth Assessment cycle embers diagrams have compared risks for different socio-economic pathways and for different adaptive capacities with sea level rise as an axis. Images published in literature outside of the IPCC have varied more widely over time but often retained similar colour schemes.

Figure 3. Comparison of risk thresholds across IPCC Assessments. All burning embers are presented with the same colour and temperature scale, removing technical details that varied between the original publications while carefully keeping the levels of the colour changes. The vertical scale indicates global mean surface temperature increase above pre-industrial. The grey areas on top of each column correspond to temperatures above the assessed range in the corresponding report. Dashed lines indicate the middle of the “risk transitions” (e.g. half-way between moderate and high risk), illustrating how it evolved with each subsequent publication^{25,26}. The graphical scale of risk categories (that is, the horizontal colour bar) appeared first in AR5; earlier reports used slightly different words to define the risk levels. See Supplementary Material for details about the production of this figure¹³⁵.

Figure 4. Flowchart of the steps involved in expert elicitation used in the SRCCL. The methods included a standardized literature review, development of assessment protocol, three rounds of expert elicitation combining elements of both the Delphi technique and SHELF method.

Figure 5. Expert elicitation for the Food Stability ember in SRCCL. The vertical axis represents GMT increase since the pre-industrial period. In panels a and b, the curves show probability density functions for each transition - undetectable to moderate in yellow; moderate to high in red; high to very high in purple. Those probability distributions were estimated on the basis of the set of expert judgements about the GMT increase corresponding to each transition, for two successive elicitation rounds. The temperature range of two of the distributions - undetectable to moderate and high to very high - ‘shrank’ between the 1st (panel a) and 2nd round (panel b). This indicates increasing consensus around transitions as experts had the chance to consider justifications and evidence presented by colleagues in the first round of elicitation. For moderate to high risk, the distribution did not change substantially, demonstrating that the elicitation does not always result in convergence. After the third round, a group discussion, the final temperature ranges for transitions identified were (panel c).

Figure 6. Flowchart of the steps involved in the sea-level rise risk assessment in the SROCC. A standardized scoring system was created to assess the relative contribution of nine metrics to risk at present and by 2100 against three sea-level rise scenarios.

References

1. Lenton, T. *et al.* Climate tipping points - too risky to bet against. *Nature* (2019).
2. United Nations. United Nations Framework Convention on Climate Change. (1992).
3. Garner, G., Reed, P. & Keller, K. Climate risk management requires explicit representation of societal trade-offs. *Clim. Change* (2016). doi:10.1007/s10584-016-1607-3
4. Managing the global commons: the economics of climate change. *Choice Rev. Online* (1995). doi:10.5860/choice.32-6292
5. Leemans, R. & Vellinga, P. The scientific motivation of the internationally agreed 'well below 2 °C' climate protection target: a historical perspective. *Curr. Opin. Environ. Sustain.* **26–27**, 134–142 (2017).
6. Fischlin, A., Ji, Z., Vladu, F. & Bisiaux, A. *Report on the Structured Expert Dialogue on the 2013–2015 Review of the United Nations Framework Convention on Climate Change (UNFCCC). Final Report FCCC/SB/2015/INF.1.* (2015).
7. Smith, J. *et al.* Vulnerability to climate change and reasons for concern: a synthesis. in *Climate Change 2001 - Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. McCarthy, J. J., Canziani, O. F., Leary, N. A., Dokken, D. J. & White, K. S.) 913–967 (Cambridge University Press, 2001).
8. *Climate Change 2001: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC).* (Cambridge University Press, 2001).
9. O'Neill, B. C. *et al.* IPCC reasons for concern regarding climate change risks. *Nat. Clim. Chang.* **7**, 28–37 (2017).
10. Matthews, J. B. R. (ed. . Annex I: Glossary. in *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change*, (eds. [Masson-Delmotte, V. et al.] (IPCC, 2018).
11. United Nations. Paris Agreement. (2015).
12. Fischlin, A., Ji, Z., Vladu, F. & Bisiaux, A. *Report on the Structured Expert Dialogue on the 2013–2015 Review of the United Nations Framework Convention on Climate Change.* (2015).
13. IPCC. *Chapter Outline of the Working Group II Contribution to the IPCC Sixth Assessment Report (AR6). 46th Session of the IPCC* (2017).
14. Mahony, M. Climate change and the geographies of objectivity: The case of the IPCC's burning embers diagram. *Trans. Inst. Br. Geogr.* (2015). doi:10.1111/tran.12064
15. Oppenheimer, M., Little, C. M. & Cooke, R. M. Expert judgement and uncertainty quantification for climate change. *Nat. Clim. Chang.* **6**, 445–451 (2016).
16. Budescu, D. V., Por, H. H. & Broomell, S. B. Effective communication of uncertainty in the IPCC reports. *Clim. Change* (2012). doi:10.1007/s10584-011-0330-3
17. Swart, R., Bernstein, L., Ha-Duong, M. & Petersen, A. Agreeing to disagree: Uncertainty management in assessing climate change, impacts and responses by the IPCC. *Climatic Change* (2009). doi:10.1007/s10584-008-9444-7
18. Socolow, R. H. High-consequence outcomes and internal disagreements: Tell us more, please. *Clim. Change* (2011). doi:10.1007/s10584-011-0187-5
19. Adler, C. E. & Hirsch Hadorn, G. The IPCC and treatment of uncertainties: Topics and sources of dissensus. *Wiley Interdisciplinary Reviews: Climate Change* (2014). doi:10.1002/wcc.297
20. Sutton, R. T. ESD Ideas: A simple proposal to improve the contribution of IPCC WGI to the assessment and communication of climate change risks. *Earth Syst. Dyn.* (2018). doi:10.5194/esd-9-1155-2018
21. Yohe, G. & Oppenheimer, M. Evaluation, characterization, and communication of

- uncertainty by the intergovernmental panel on climate change-an introductory essay. *Clim. Change* (2011). doi:10.1007/s10584-011-0176-8
22. Moss, R. H. & Schneider, S. H. Uncertainties in the IPCC TAR: Recommendations to lead authors for more consistent assessment and reporting. in *Moss, R. H. & Schneider, S. H., 2000. Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC. Intergovernmental Panel on Climate Change (IPCC)* (eds. Pachauri, R., Taniguchi, T. & Tanaka, K.) 33–51 (World Meteorological Organisation (WMO)), 2000).
 23. Manning, M. *et al.* *Describing scientific uncertainties in climate change to support analysis of risk and of options.* (2004).
 24. Mastrandrea, M. D. *et al.* *Guidance note for lead authors of the IPCC Fifth Assessment Report on consistent treatment of uncertainties.* (2010). doi:<https://www.ipcc.ch/site/assets/uploads/2018/05/uncertainty-guidance-note.pdf>
 25. Jones, J. & Hunter, D. Qualitative Research: Consensus methods for medical and health services research. *BMJ* (1995). doi:10.1136/bmj.311.7001.376
 26. Peel, A. *et al.* Use of Expert Judgement Across NICE Guidance-Making Programmes: A Review of Current Processes and Suitability of Existing Tools to Support the Use of Expert Elicitation. *Appl. Health Econ. Health Policy* (2018). doi:10.1007/s40258-018-0415-5
 27. Waggoner, J., Carline, J. D. & Durning, S. J. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. *Academic Medicine* (2016). doi:10.1097/ACM.0000000000001092
 28. Butler, A. J., Thomas, M. K. & Pintar, K. D. M. Systematic review of expert elicitation methods as a tool for source attribution of enteric illness. *Foodborne Pathogens and Disease* (2015). doi:10.1089/fpd.2014.1844
 29. Morgan, M. G. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences of the United States of America* (2014). doi:10.1073/pnas.1319946111
 30. Aspinall, W. P. & Cooke, R. M. *Quantifying scientific uncertainty from expert judgement elicitation. Risk and Uncertainty Assessment for Natural Hazards* **9781107006**, (2011).
 31. Graefe, A. & Armstrong, J. S. Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task. *Int. J. Forecast.* (2011). doi:10.1016/j.ijforecast.2010.05.004
 32. Murphy *et al.* Consensus development methods, and their use in clinical guideline development. *Health Technol. Assess. (Rockv)*. (1998). doi:10.3310/hta2030
 33. IPCC. IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse gas fluxes in Terrestrial Ecosystems. *Summ. Policymakers Approv. Draft* (2019). doi:10.4337/9781784710644
 34. IPCC. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. in *IPCC Summary for Policymakers* (2019). doi:<https://www.ipcc.ch/report/srocc/>
 35. Tol, R. S. J. Equitable cost-benefit analysis of climate change policies. *Ecol. Econ.* (2001). doi:10.1016/S0921-8009(00)00204-4
 36. Tol, R. S. J. Estimates of the damage costs of climate change: Part 1: Benchmark estimates. *Environ. Resour. Econ.* (2002). doi:10.1023/A:1014500930521
 37. Tol, R. S. J. Estimates of the damage costs of climate change: Part II. Dynamic estimates. *Environ. Resour. Econ.* (2002). doi:10.1023/A:1014539414591
 38. Smith, J. *et al.* First Order Draft: Vulnerability to climatechange and reasons for concern. in *Climate Change 2001 - Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel of Climate Change* (1999).
 39. Smith, J. *et al.* Second Order Draft: Vulnerability to Climate Change and Reasons for Concern. in *Climate Change 2001 - Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel of Climate Change* (2000).
 40. Smith, J. personal communication.

41. *Climate change 2001: Synthesis report - A Report of the Intergovernmental Panel on Climate Change (IPCC). Stand-alone edition. Climate change 2001: Synthesis report* (Cambridge University Press, 2001).
42. Caddy, J. F. Limit reference points, traffic lights, and holistic approaches to fisheries management with minimal stock assessment input. *Fisheries Research* (2002). doi:10.1016/S0165-7836(01)00343-5
43. Wilson, D. C. & Pascoe, S. Chapter 13 Delivering complex scientific advice to multiple stakeholders. *Dev. Aquac. Fish. Sci.* (2006). doi:10.1016/S0167-9309(06)80016-5
44. Bystrom, A. B., Naranjo-Madrigal, H. & Wehrtmann, I. Indicator-based management recommendations for an artisanal bottom-longline fishery in Costa Rica, Central America. *Rev. Biol. Trop.* (2017). doi:10.15517/rbt.v65i2.25757
45. Emrich, T. E., Qi, Y., Lou, W. Y. & L'Abbe, M. R. Traffic-light labels could reduce population intakes of calories, total fat, saturated fat, and sodium. *PLoS One* (2017). doi:10.1371/journal.pone.0171188
46. Kunz, S., Haasova, S., Rieß, J. & Florack, A. Beyond healthiness: The impact of traffic light labels on taste expectations and purchase intentions. *Foods* (2020). doi:10.3390/foods9020134
47. Simms, R. A. *et al.* Development of maternity dashboards across a UK health region; Current practice, continuing problems. *Eur. J. Obstet. Gynecol. Reprod. Biol.* (2013). doi:10.1016/j.ejogrb.2013.06.003
48. Donnelly, L. & Harrison, M. Geomorphological and geoforensic interpretation of maps, aerial imagery, conditions of diggability and the colour-coded RAG prioritization system in searches for criminal burials. *Geol. Soc. Spec. Publ.* (2013). doi:10.1144/SP384.10
49. Ruffell, A. A RAG System for the Management Forensic and Archaeological Searches of Burial Grounds. *Int. J. Archaeol.* (2015). doi:10.11648/j.ija.s.2015030101.11
50. Boyle, I. & Toft, M. Traffic Light Decision Making A Simple Model To Evaluate Dangers And Guide Safe Decision Making In Outdoor Adventure Experiences. *J. Outdoor Environ. Educ.* (1998). doi:10.1007/bf03400694
51. Mahony, M. & Hulme, M. The Colour of Risk: An Exploration of the IPCC's "Burning Embers" Diagram. *Spontaneous Gener. A J. Hist. Philos. Sci.* (2012). doi:10.4245/sponge.v6i1.16075
52. Vellinga, P. & Swart, R. The greenhouse marathon: proposal for a global strategy. *Clim. Chang. Proc. 2nd World Clim. Conf. Geneva, 1990* (1991).
53. Vellinga, P. & Swart, R. The greenhouse marathon: A proposal for a global strategy - A guest editorial. *Climatic Change* (1991). doi:10.1007/BF00142501
54. Leemans, R. & Vellinga, P. The scientific motivation of the internationally agreed 'well below 2 °C' climate protection target: a historical perspective. *Current Opinion in Environmental Sustainability* (2017). doi:10.1016/j.cosust.2017.07.010
55. Schneider, B. Burning worlds of cartography: a critical approach to climate cosmograms of the Anthropocene. *Geo Geogr. Environ.* (2016). doi:10.1002/geo2.27
56. White, K. S. *et al.* Technical Summary Climate Change 2001: Impacts, Adaptation and Vulnerability. in *Climate Change 2001 - Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel of Climate Change* (eds. McCarthy, J. J., Canziani, O. F., Leary, N. A. & Dokken, D. J.) (IPCC, 2001).
57. Mastrandrea, M. D. & Schneider, S. H. Probabilistic Integrated Assessment of 'Dangerous' Climate Change. *Science* (80-.). (2004). doi:10.1126/science.1094147
58. Gattuso, J. P. *et al.* Contrasting futures for ocean and society from different anthropogenic CO₂ emissions scenarios. *Science* (2015). doi:10.1126/science.aac4722
59. Magnan, A. K. *et al.* Implications of the Paris agreement for the ocean. *Nat. Clim. Chang.* **6**, 732–735 (2016).
60. Fischlin, A. Berücksichtigen wir in der Klimapolitik genügend Sicherheitsmargen? Do We Have Sufficient Safety Margins in Climate Policy? *GAIA - Ecol. Perspect. Sci. Soc.* (2009). doi:10.14512/gaia.18.3.3

61. Schneider, S. H. *et al.* Assessing key vulnerabilities and the risk from climate change. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 779–810 (Cambridge University Press, 2007).
62. Smith, J. B. *et al.* Assessing dangerous climate change through an update of the Intergovernmental Panel on Climate Change (IPCC) “reasons for concern”. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0812355106
63. IPCC. Summary for policymakers. in *Climate change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)* (eds. Solomon, S. et al.) 1–18 (Cambridge University Press, 2007).
64. IPCC. Summary for policymakers. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 7–22 (Cambridge University Press, 2007).
65. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E. Cross-chapter case studies. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 843–868 (Cambridge University Press, 2007).
66. Fischlin, A. *et al.* Ecosystems, their properties, goods and services. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 211–272 (Cambridge University Press, 2007).
doi:http://www.ipcc.ch/publications_and_data/ar4/wg2/en/ch4.html
67. Hennessy, K. *et al.* Australia and New Zealand. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 507–540 (Cambridge University Press, 2007).
68. Technical summary. in *Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change (IPCC)* (eds. Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J. & Hanson, C. E.) 23–78 (Cambridge University Press, 2007).
69. Oppenheimer, M. *et al.* Emergent risks and key vulnerabilities. in *Climate change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)* (eds. Field, C. B. et al.) A, 1039–1099 (Cambridge University Press, 2014).
70. IPCC *et al.* Summary for Policymakers. in *Climate change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)* (eds. Field, C. B. et al.) A, In press (IPCC with World Meteorological Organisation (WMO), and United Nations Environmental Program (UNEP), 2013).
71. Field, C. B. *et al.* *Managing the risks of extreme events and disasters to advance climate change adaptation: Special report of the intergovernmental panel on climate change. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change* (2012). doi:10.1017/CBO9781139177245
72. IPCC *et al.* *Managing the Risks of Extreme Events and Disasters to Advance Climate Change*

- Adaptation - SREX Summary for Policymakers. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change.* (2012).
doi:10.1017/CBO9781139177245
73. Hoegh-Guldberg, O. *et al.* Impacts of 1.5°C global warming on natural and human systems. in *Global Warming of 1.5 °C an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change* (2018).
 74. Palazzo, A. *et al.* Linking regional stakeholder scenarios and shared socioeconomic pathways: Quantified West African food and climate futures in a global context. *Glob. Environ. Chang.* **45**, 227–242 (2017).
 75. Byers, E. *et al.* Global exposure and vulnerability to multi-sector development and climate change hotspots. *Environ. Res. Lett.* **13**, 055012 (2018).
 76. Hanasaki, N. *et al.* A global water scarcity assessment under Shared Socio-economic Pathways - Part 2: Water availability and scarcity. *Hydrol. Earth Syst. Sci.* (2013). doi:10.5194/hess-17-2393-2013
 77. Hanasaki, N. *et al.* A global water scarcity assessment under Shared Socio-economic Pathways - Part 1: Water use. *Hydrol. Earth Syst. Sci.* (2013). doi:10.5194/hess-17-2375-2013
 78. Oppenheimer, M. *et al.* Sea Level Rise and Implications for Low Lying Islands, Coasts and Communities. in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (eds. Pörtner, H.-O. *et al.*) (2019).
 79. World Meteorological Organization. *United In Science: High-level synthesis report of latest climate science information convened by the Science Advisory Group of the UN Climate Action Summit 2019.* (2019).
 80. Robinson, A., Calov, R. & Ganopolski, A. Multistability and critical thresholds of the Greenland ice sheet. *Nat. Clim. Chang.* (2012). doi:10.1038/nclimate1449
 81. Levermann, A. *et al.* The multimillennial sea-level commitment of global warming. *Proc. Natl. Acad. Sci. U. S. A.* (2013). doi:10.1073/pnas.1219414110
 82. Levermann, A. *et al.* Projecting Antarctic ice discharge using response functions from SeaRISE ice-sheet models. *Earth Syst. Dyn.* (2014). doi:10.5194/esd-5-271-2014
 83. Golledge, N. R. *et al.* The multi-millennial Antarctic commitment to future sea-level rise. *Nature* (2015). doi:10.1038/nature15706
 84. Fürst, J. J., Goelzer, H. & Huybrechts, P. Ice-dynamic projections of the Greenland ice sheet in response to atmospheric and oceanic warming. *Cryosphere* (2015). doi:10.5194/tc-9-1039-2015
 85. Pattyn, F. *et al.* The Greenland and Antarctic ice sheets under 1.5 °C global warming. *Nature Climate Change* (2018). doi:10.1038/s41558-018-0305-8
 86. Hoegh-Guldberg, O. *et al.* Special Report on Global Warming of 1.5 °C - Chapter 3: Impacts of 1.5° C global warming on natural and human systems. *Glob. Warm. 1.5°C. An IPCC Spec. Rep. impacts Glob. Warm. 1.5°C above pre-industrial levels Relat. Glob. Greenh. gas Emiss. pathways, Context Strength. Glob. response to Threat Clim. Chang.* (2018). doi:10.1002/ejoc.201200111
 87. Pörtner, H.-O. *et al.* Ocean systems. in *Climate change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)* (eds. Field, C. B. *et al.*) **A**, 411–484 (Cambridge University Press, 2014).
 88. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. *Process and methods guides* (2014).
 89. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. *Natl. Inst. Heal. Care Excell.* (2013). doi:10.2165/00019053-200826090-00002
 90. Stafford, A. C., Bindoff, I. K., Tenni, P. C., Peterson, G. M. & Doran, C. M. A methodological framework for estimating the clinical and economic value of community pharmacists' clinical interventions using expert opinion. *Journal of Clinical Pharmacy and Therapeutics*

- (2012). doi:10.1111/j.1365-2710.2011.01322.x
91. Khodyakov, D. *et al.* Conducting Online Expert panels: A feasibility and experimental replicability study. *BMC Med. Res. Methodol.* (2011). doi:10.1186/1471-2288-11-174
 92. Sinha, I. P., Smyth, R. L. & Williamson, P. R. Using the Delphi technique to determine which outcomes to measure in clinical trials: Recommendations for the future based on a systematic review of existing studies. *PLoS Med.* (2011). doi:10.1371/journal.pmed.1000393
 93. Moher, D., Schulz, K. F., Simera, I. & Altman, D. G. Guidance for developers of health research reporting guidelines. *PLoS Medicine* (2010). doi:10.1371/journal.pmed.1000217
 94. Ferri, C. P. *et al.* Global prevalence of dementia: A Delphi consensus study. *Lancet* (2005). doi:10.1016/S0140-6736(05)67889-0
 95. Cresswell, K. M. *et al.* Global Research Priorities to Better Understand the Burden of Iatrogenic Harm in Primary Care: An International Delphi Exercise. *PLoS Med.* (2013). doi:10.1371/journal.pmed.1001554
 96. Hasson, F., Keeney, S. & McKenna, H. Research guidelines for the Delphi survey technique. *J. Adv. Nurs.* (2000). doi:10.1046/j.1365-2648.2000.t01-1-01567.x
 97. Shrader-Frechette, K. Experts in Uncertainty: Opinion and Subjective Probability in Science. Roger M. Cooke. *Ethics* (1993). doi:10.1086/293541
 98. Ferguson, J. H. The NIH Consensus Development Program: The evolution of guidelines. *Int. J. Technol. Assess. Health Care* (1996). doi:10.1017/S0266462300009818
 99. Dalkey, N. & Helmer, O. An Experimental Application of the DELPHI Method to the Use of Experts. *Manage. Sci.* (1963). doi:10.1287/mnsc.9.3.458
 100. Delbecq, A. L. & Van de Ven, A. H. A Group Process Model for Problem Identification and Program Planning. *J. Appl. Behav. Sci.* (1971). doi:10.1177/002188637100700404
 101. Dias, L. C., Morton, A. & Quigley, J. Elicitation: State of the art and science. in *International Series in Operations Research and Management Science* (2018). doi:10.1007/978-3-319-65052-4_1
 102. Black, N. *et al.* Consensus development methods: A review of best practice in creating clinical guidelines. *J. Heal. Serv. Res. Policy* (1999). doi:10.1177/135581969900400410
 103. James, D. & Warren-Forward, H. Research methods for formal consensus development. *Nurse Res.* (2015). doi:10.7748/nr.22.3.35.e1297
 104. Cooke, R., Mendel, M. & Thijs, W. Calibration and information in expert resolution; a classical approach. *Automatica* (1988). doi:10.1016/0005-1098(88)90011-8
 105. Colson, A. R. & Cooke, R. M. Expert elicitation: Using the classical model to validate experts' judgments. *Rev. Environ. Econ. Policy* (2018). doi:10.1093/leep/rex022
 106. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. *EFSA J.* (2014). doi:10.2903/j.efsa.2014.3734
 107. Clemen, R. T. & Winkler, R. L. Combining probability distributions from experts in risk analysis. *Risk Anal.* (1999). doi:10.1023/A:1006917509560
 108. Walker, K. D., Evans, J. S. & MacIntosh, D. Use of expert judgment in exposure assessment part I. Characterization of personal exposure to benzene. *J. Expo. Anal. Environ. Epidemiol.* (2001). doi:10.1038/sj.jea.7500171
 109. McCormack, C. G. *et al.* Key impacts of climate engineering on biodiversity and ecosystems, with priorities for future research. *J. Integr. Environ. Sci.* (2016). doi:10.1080/1943815X.2016.1159578
 110. Griscom, B. W. *et al.* Natural climate solutions. *Proc. Natl. Acad. Sci. U. S. A.* (2017). doi:10.1073/pnas.1710465114
 111. Mach, K. J. *et al.* Climate as a risk factor for armed conflict. *Nature* (2019). doi:10.1038/s41586-019-1300-6
 112. Bamber, J. L. & Aspinall, W. P. An expert judgement assessment of future sea level rise from the ice sheets. *Nat. Clim. Chang.* (2013). doi:10.1038/nclimate1778
 113. Mukherjee, N. *et al.* The Delphi technique in ecology and biological conservation: Applications and guidelines. *Methods Ecol. Evol.* **6**, 1097–1109 (2015).

114. Bhave, A. G., Conway, D., Dessai, S. & Stainforth, D. A. Water Resource Planning Under Future Climate and Socioeconomic Uncertainty in the Cauvery River Basin in Karnataka, India. *Water Resour. Res.* (2018). doi:10.1002/2017WR020970
115. Zommers, Z. & Alverson, K. Intelligent tinkering in climate change adaptation. in *Resilience: The Science of Adaptation to Climate Change* (2018). doi:10.1016/B978-0-12-811891-7.00027-X
116. Berrang-Ford, L., Pearce, T. & Ford, J. D. Systematic review approaches for climate change adaptation research. *Regional Environmental Change* (2015). doi:10.1007/s10113-014-0708-7
117. Stanton, C. Y. *et al.* Managing cropland and rangeland for climate mitigation: an expert elicitation on soil carbon in California. *Clim. Change* (2018). doi:10.1007/s10584-018-2142-1
118. Zickfeld, K., Morgan, M. G., Frame, D. J. & Keith, D. W. Expert judgments about transient climate response to alternative future trajectories of radiative forcing. *Proc. Natl. Acad. Sci. U. S. A.* (2010). doi:10.1073/pnas.0908906107
119. Kriegler, E., Hall, J. W., Held, H., Dawson, R. & Schellnhuber, H. J. Imprecise probability assessment of tipping points in the climate system. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0809117106
120. Hurlbert, M. & Krishnaswamy, E. Davin, F.X. Johnson, C.F. Mena, J. Morton, S. Myeong, D. Viner, K. Warner, A. Wreford, S. Zakieldean, Zommers, Z. Climate Change and Land: Risk management and decision-making in relation to sustainable development. *IPCC Spec. Rep.* (2019).
121. Bantel, K. A. Comprehensiveness of Strategic Planning: The Importance of Heterogeneity of a Top Team. *Psychol. Rep.* (1993). doi:10.2466/pr0.1993.73.1.35
122. Smith, P., J. Nkem, K. Calvin, D. Campbell, F. Cherubini, G. Grassi, V. Korotkov, A.L. Hoang, S. Lwasa, P. McElwee, E. Nkonya, N. Saigusa, J.-F. Soussana, M. A. T. *Climate Change and Land: Interlinkages between desertification, land degradation, food security and greenhouse gas fluxes: Synergies, trade-offs and integrated response options. IPCC Special Report* (2019).
123. Bolger, F. & Wright, G. Improving the Delphi process: Lessons from social psychological research. *Technol. Forecast. Soc. Change* (2011). doi:10.1016/j.techfore.2011.07.007
124. Jairath, N. & Weinstein, J. The Delphi methodology (Part one): A useful administrative approach. *Can. J. Nurs. Adm.* (1994).
125. Bolger, F., Stranieri, A., Wright, G. & Yearwood, J. Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters? *Technol. Forecast. Soc. Change* (2011). doi:10.1016/j.techfore.2011.06.002
126. Singh, G. G. *et al.* Group elicitation yield more consistent, yet more uncertain experts in understanding risks to ecosystem services in New Zealand bays. *PLoS One* (2017). doi:10.1371/journal.pone.0182233
127. Speirs-Bridge, A. *et al.* Reducing overconfidence in the interval judgments of experts. *Risk Anal.* (2010). doi:10.1111/j.1539-6924.2009.01337.x
128. Sperber, D., Mortimer, D., Lorgelly, P. & Berlowitz, D. An expert on every street corner? Methods for eliciting distributions in geographically dispersed opinion pools. *Value Heal.* **16**, 434–437 (2013).
129. Adger, W. N., Brown, I. & Surminski, S. Advances in risk assessment for climate change adaptation policy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (2018). doi:10.1098/rsta.2018.0106
130. Yohe, G. 'Reasons for concern' (about climate change) in the United States. *Climatic Change* (2010). doi:10.1007/s10584-009-9797-6
131. Warszawski, L. *et al.* The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proceedings of the National Academy of Sciences of the United States of America* (2014). doi:10.1073/pnas.1312330110
132. Frieler, K. *et al.* Assessing the impacts of 1.5°C global warming - Simulation protocol of

- the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geosci. Model Dev.* (2017). doi:10.5194/gmd-10-4321-2017
133. Sharpe, S. Telling the boiling frog what he needs to know: why climate change risks should be plotted as probability over time. *Geosci. Commun. Discuss.* (2019). doi:10.5194/gc-2019-1
 134. IPCC. Climate Change 2014 Synthesis Report Summary Chapter for Policymakers. *Ippc* (2014). doi:10.1017/CBO9781107415324
 135. Marbaix, P. The Embers Facotory. (2020). Available at: <https://www.elic.ucl.ac.be/users/marbaix/ember-factory> .

Supplementary Information

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements and disclaimer

The authors would like to thank Joel Smith for useful reviews and informed comments on the paper. A.K.M. thanks the French National Research Agency (“Investment for the Future” programme, ANR-10-LABX-14-01). All views expressed herein are the responsibility of the authors and do not necessarily reflect their respective institutions, the IPCC or the United Nations.

Author Contributions

Z.Z. and W.T. designed the study. Z.Z., E.D., W.T., K.C., M.H., K.W., J.P.E., and S.G. developed the improved expert elicitation procedure for SRCCL. A.K.M., H.-O.P., and Z.S contributed to the SROCC burning embers development. P.M. harmonised the burning embers for three RFC across the IPCC reports and developed Figure 2, 3. M.H., P.M. and W.T lead analysis and design of Figure 5 . Z.I., A.F. and C.R. contributed substantial additions to the history section. B.O. and P.M. helped research and draft the section on changes in risks. S.G., M.H., Z.I wrote the section on expert elicitation. Z.Z. lead the drafting of the paper. All authors contributed to the editing and revision of the manuscript and helped develop conclusions.

Author Information

Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests.