

'Telling the Boiling Frog What He Needs to Know'

Review and discussion summary: comments, responses and changes to manuscript

Comment 1a

5 I agree with all that is proposed in this piece, and I think it is important to try and persuade the assessment community to adopt multiple perspectives and multiple communication devices to increase the cogency and relevance of climate change communication. This proposal is one, and a valuable one, and I'd be happy to see this piece published. Mine are not really requests for specific corrections or changes, I just want to point out a few things. In AR6 WG1 there is going to be a new chapter, chapter 12, which is attempting to facilitate exactly the type of cross-WG efforts the author is highlighting and wishing
10 for in this piece. The authors of Chapter 12 (I'm one) are trying to identify metrics and thresholds that have impact relevance, and whose changes can be assessed from observations and climate model output from the point of view of WG1 science (i.e. remaining strictly within the boundaries of hazard characterization). It is far from easy, and that may be perhaps better acknowledged in this piece. I think the type of approach that is called for here makes a lot of sense for specific impact analyses, but it is far from straightforward in big, global assessments (or even national ones) given the myriad of metrics/thresholds that
15 are relevant to some sectors/regions and less relevant to others. But an attempt to at least identify some examples, in a multidisciplinary approach, is being made. So hopefully next time around some parts of the IPCC report will have more immediate resonance with what is being wished for here.

Response 1a

It's wonderful to know that work is going on to identify relevant impact thresholds for discussion in AR6. I agree identifying
20 meaningful thresholds is not always easy. I think in most cases it takes a fair amount of expert judgment – for example, even in relation to a biophysical threshold such as the limit of human tolerance for heat stress, a judgment has to be made about how many hours of heat and humidity conditions above that limit is to be defined as an instance of 'crossing the threshold', around which estimates of probability (as a function of time) can be made. I'm happy to give that difficulty more acknowledgement. I suppose one way of saying this is that in almost no case is there a single 'correct' definition of the threshold, and this makes
25 it difficult to choose the one that's most appropriate. At the same time, I'd argue that almost any reasonably defensible choice of threshold is better than no choice at all.

I also agree that many thresholds will be specific to their location, and so not applicable at a national or global scale. A good example of this was once recounted to me by a very experienced climate scientist in the US. After I had outlined to him my argument for taking a probability over time approach, he recalled one time his research team had done that. They had asked a
30 city government what level of extreme rainfall they were worried about. The city government said it was X inches of rainfall over a period of Y hours (I don't recall the numbers, but they gave them), because this was the level that caused the city's sewer system to overflow, bringing faeces floating onto the streets. (That was their non-arbitrary threshold). The research team went away and came back with a plot of the probability of faeces floating in the streets as a function of time. The scientist

told me that in all his years of advising this city government, he could not think of any other time when his advice had made such a strong impression. It was a great example of a non-arbitrary threshold, but perhaps not a widely replicable one.

Despite the local specificity of many thresholds, I think a collection of threshold-based risk assessments could add real value to either a national or a global assessment. For example, knowing that several parts of the world could exceed human tolerance for heat stress, that many places could exceed the temperature tolerances for several important crops, and that other places could fall below socioeconomic thresholds of minimum renewable water resources, can greatly enrich our understanding of the scale of the risks facing the world as a whole. There is no need for this information to be aggregated into global-scale metrics, or even for the same thresholds to be used for different regions, for this information to be useful. There may be an issue here that is beyond the scope of this paper, but I wonder whether there is a difference between the value of consistency of measurement for the purposes of advancing scientific understanding, and the value of diversity of measurement for the purposes of informing a risk assessment.

The main difficulty I've mentioned in the initial draft of this paper is the need to start – before doing any science – with a subjective question: 'what is it that we wish to avoid?' The point I'm making here is about the *order* in which things need to be done. In the heat stress example mentioned in the paper (and presented in this report <http://www.csap.cam.ac.uk/projects/climate-change-risk-assessment/>) the order was: first, have a policymaker define an issue of concern (environmental conditions exceeding human tolerance for heat stress); second, have an expert in the system of concern (human health) from the WG2 community define a meaningful threshold in relation to that system (X hours above Y degrees WBGT); and last, have an expert in the physical climate from the WG1 community estimate the probability of crossing the threshold as a function of time, for some regions where it looks likely to be a problem. Done this way round, it's quite feasible, but if you start with the WG1 science, it's almost impossible. I'd be interested to know whether you agree.

Change 1a

New text, p7 lines 7-11: The probability over time approach is not without its challenges. Most thresholds do not have a single 'correct' definition. For example, a temperature tolerance threshold may be biophysical, but how long above that temperature should be considered 'too long' is a matter of expert judgment. Many thresholds are specific to their location, meaning that a national or global risk assessment performed this way may need to be made up of a diverse range of quite distinct studies, rather than using the kind of consistent data that allows for aggregation.

Comment 1b

One thing I would like to see discussed in a more nuanced way is the idea of global warming levels and targets. I appreciate the recognition that these are sociopolitical constructs: it is in my opinion an important distinction that needs to be made. I agree they have their "aspirational value" and as such they may be taken as the lens through which to carry out this type of communication, but I would like to see an explicit mention of the fact that there is no such thing as an absolute threshold below which we are safe and above which we are toast, at this global scale, and that makes the treatment of such thresholds very different from those that have a physical meaning. Gavin Schmidt had a good piece, a long time ago on RealClimate about

this. <http://www.realclimate.org/index.php/archives/2006/07/runaway-tipping-points-of-no-return/> I would be happy if this disclaimer was made explicit here, since at some level I feel as if focusing on this communication for such arbitrary threshold may sometime be less than productive.

Response 1b

5 Regarding global warming thresholds, I'm happy to make an explicit mention of the fact that, as you say, 'there is no such thing as an absolute threshold below which we are safe and above which we are toast'. This is clearly true. I think thresholds with a physical basis are likely to be the most dangerous to cross, and so are particularly important to investigate. Socioeconomic thresholds may perhaps be more amenable to adaptation. What I've called 'experiential' thresholds may have only transient value – they are meaningful only for as long as people have strong memories of the original event. I think it
10 would be going too far to attempt to rank different kinds of thresholds by how 'non-arbitrary' they are, but I agree it is important to distinguish between them, and useful to recognise their different characteristics.

The Gavin Schmidt piece is very good, and makes the point well that there is not just one tipping point in the climate system. It also illustrates the difficulty of defining meaningful non-arbitrary thresholds in relation to changes in the climate system itself (strictly within WG1 scope), as opposed to thresholds in relation to impacts on people (WG2 scope, to which everything
15 I've said above refers). However, it seems that even the relatively feasible examples are not always communicated as clearly as they could be. Schmidt gives the disappearance of the Greenland Ice Sheet as one of the best examples of a 'point of no return'. The IPCC's Special Report on 1.5°C did helpfully mention that 'these instabilities could be triggered at around 1.5°C to 2°C of global warming', but gave most prominence to the impact-over-time finding that the most likely global sea level rise at 2100 would be around 10cm lower at 1.5°C than at 2°C. It might have been more helpful to compare the *probability* of
20 triggering disappearance of the GIS at (a sustained) 1.5°C, compared to that probability at 2°C. If illustrated by data such as those underlying figure 13.14c in AR5 WG1, this could perhaps have presented a very powerful picture of the difference in risk between the two levels of warming.

Change 1b

New text, p3 lines 25-26: There is no temperature threshold within which we are safe, and beyond which we are all cooked.

25 New text, p6 lines 26-30: Clearly, these different kinds of threshold vary in their objectivity, and in other ways likely to have implications for risk assessment. Socioeconomic thresholds may be more possible to overcome through adaptation than those defined by physical or biochemical properties alone. Experiential and political thresholds may have only transient value – useful for as long as the past event is remembered, or for as long as the policy holds. But what all these thresholds have in common is that they are relevant to policymakers, because they are defined in terms of what it is that we collectively wish to
30 avoid.

Comment 1c

Lastly, just one point about the (idealized) examples of graphics presented in the paper. It may be good to point out that the type of deterministic relation shown by these single lines is very rarely (if ever) what we have, but that some fuzziness, or

shading, communicating uncertainty will also be part of this type of display. I know it is obvious to many of us but all the examples happen to be devoid of uncertainty, and that is possibly misleading for a reader.

Response 1c

5 Finally, regarding shaded uncertainty bands, I agree this is worth mentioning. I thought I had done so, but hadn't. I'll add this in. A spectacular example of shaded bands showing the range of risk is figure 12.5 from AR5 WG1, which shows the uncertainty around long-term global temperature increase. Within the set of impact over time graphs, I have found these two (AR5 WG1 12.5, and 13.14 as mentioned above) exceptionally useful for communicating the extent of the risks, but have never met another non-scientist who was aware of them – probably because they were not included in the SPM. I can only guess they were excluded from the SPM because they were considered too long-term to be policy relevant, but arguably, since 10 they show the largest risks, they are the most policy relevant of all. I think including long-enough time horizons for the largest risks to be visible in all SPMs could be helpful – and would be consistent with the general principle that a risk assessment should always consider the biggest risks.

Change 1c

15 New text, p2 lines 18-20: (Shaded bands illustrating uncertainty in impact can bring a broader range of risks into view, but still provide no guarantee that the biggest risks will be visible.)

Comment 2a

This discussion piece raises some important issues. I'd be happy to see it published. I don't have any specific corrections or requested changes but I'll make a few comments.

20 The article argues that research should start by identifying what it is that we most want to avoid and then go on to assess the likelihood of this as a function of time. The author suggests that research funders have a key role to play in this endeavour “by structuring research calls in a way that mandates the inter-disciplinary approach and the pre-research engagement with decision-makers”. I see significant merit in adopting measures of this type and I support the author in these goals. Nevertheless there are risks in such an approach and I think the article would benefit from more nuance on certain points.

25 It is certainly true that left to its own devices climate change science can focus on issues that are more removed from societal relevance than they should be. However, the article seems to argue that climate change science has a simplistic role of answering the questions asked by society, as if it is simply a matter of turning a prediction handle. Pre-research engagement with decision makers is something I agree would have significant value but by structuring calls towards answering specific decision-relevant questions one can inadvertently encourage research which makes whatever assumptions are necessary to 30 output an answer. What can get lost in this process is reflexion on whether the question asked is currently answerable with any reasonable degree of confidence. Are the assumptions justified? Researchers or research disciplines that perceive the need for more fundamental research in order to answer the question simply don't apply because they can't address the terms of the call; as a result the outputs can become biased and over confident. The approach treats research as simply an extended form of consultancy. It risks undermining consideration of which questions can currently be answered.

Although I support most aspects of the recommendations in the paper they are not a panacea and if not addressed carefully could do more harm than good. Substantial rebalancing towards the approaches proposed would be extremely valuable but this article would be even more useful if the difficulties and risks were more thoroughly explored.

Response 2a

5 I realise I should have made clear, and will do in amending the paper, that I'm not proposing that all climate science should be done this way. Clearly, if John Tyndall hadn't been interested in the molecular physics of radiant heat, and if a century's-worth of scientists after him hadn't pursued fundamental research into the physics of the climate and all the workings of the Earth's systems, then policymakers wouldn't have any questions about climate change for any scientists to answer. I have a deep respect for fundamental research, and recognise it as the foundation that makes any policy-related inquiries possible. It's
10 right that most science should be science-led.

I am in complete agreement with the reviewer's conclusion that a 'substantial rebalancing' towards the approach I'm advocating – not a complete conversion – is what would be appropriate and valuable. My concern is that within the vast body of research being done on climate science, not enough is being done with the express intent of informing risk assessment. To give an example: when I served on the UK's delegation to approve the Summary for Policymakers (SPM) of the AR5 Working
15 Group 2 report on Impacts, I was surprised that mention of research finding that heat and humidity conditions could potentially exceed the limits of human physiological tolerance for heat stress was not being included in the SPM. I asked the lead authors whether they doubted the validity of this research finding. They said no, and that if anything, it probably understated the risk – but the rule was that a finding could only be mentioned in the SPM if it was supported by at least two separate pieces of research. And of the more than 12,000 scientific papers cited in the WG2 report, only one had asked whether a warming planet
20 might at some time, in some places, become too hot for people. (Meanwhile, I found WG2 cited nine studies that looked at the impacts of climate change on ski resorts in Europe, and thirteen studies that considered the impacts on European grape growing and wine production.)

I agree that even when research is done for the purpose of informing risk assessment, it is unlikely to be a simple process of scientists answering policymakers' questions. It is more likely to need a dialogue to arrive at a shared understanding of what
25 is possible, knowable and relevant to people's interests.

Change 2a

New text, p8 lines 8-10: The proposal is certainly not that all climate science should be done in this way. Fundamental, policy-neutral research is indispensable; and there are many ways of communicating risks. The suggestion is that more research could be done expressly for the purpose of risk assessment than is done at present, and that a deliberate approach should be taken to
30 identifying and assessing the biggest risks.

Comment 2b

A related issue is the "fuzziness" referred to in Dr. Tebaldi's review. I strongly agree that such uncertainties are a critical part of communication but it is important that these uncertainties receive the consideration they require. By demanding that climate

science answers specific questions in terms of time dependent probabilities the author is encouraging a situation where computer model ensembles are interpreted as providing such information. Maybe he has other ideas of how it would be provided but this is the most common and simplest approach. I have concerns that the approach recommended provides no pressure to robustly consider the reliability of such uncertainty assessments because this is difficult, time consuming and unlikely to be prioritised in a funder's calls if they are principally driven by the need to answer specific questions.

Response 2b

The point about uncertainty is really important. It's fair to say that the simplest approach to estimating probabilities of specific climate impacts will be through the use of computer model ensembles, and there are undoubtedly risks to this. The graphs of global temperature increase shown in Figures 2 and 3 of this paper were produced with models that did not incorporate earth system feedbacks; other work by the same authors [1] has shown that when these feedbacks are incorporated, the estimated probability of a given degree of temperature increase is substantially higher, within a wide range of uncertainty. As the actuaries wrote in our climate change risk assessment report [2], models are imperfect, and a factor that is important in determining risk should never be excluded from consideration simply because it cannot be quantified.

Having said that, I don't think the proposed approach – of assessing the probability of crossing a non-arbitrary threshold of impact as a function of time – necessarily implies using a model, or even making quantified estimates. The IPCC AR5 WG1 Chapter 12 on 'long-term climate change: projections, commitments and irreversibility' included a table [3] showing the estimated likelihood of specific abrupt and irreversible changes in the climate system occurring during the 21st century. These likelihoods were estimated using expert judgment, and were expressed in terms such as 'very unlikely', 'possible', and 'likely' instead of being quantified. The table would have been consistent with the approach I'm proposing if it had indicated how these likelihoods could change as a function of time or global temperature increase. To the extent that such a time-dependent assessment could have been supported by the science, it would have been useful.

Unquantified and judgment-based estimates of probability are common in other areas of risk assessment. A good example is counter-terrorism, where the UK government uses five 'threat levels' to indicate the likelihood of a terrorist attack in the UK [4]. These are:

- Low - an attack is unlikely
- Moderate - an attack is possible but not likely
- Substantial - an attack is a strong possibility
- Severe - an attack is highly likely
- Critical - an attack is expected imminently

In situations where there may be great danger, expert assessments such as these can be useful despite their deep uncertainties. I think the important thing is to use the best available information (whether quantified or not) to give the fullest possible assessment of the risk.

What I think this means for the recommendation about research funding is that it should be driven not so much by the need to answer specific questions, but by the need to rigorously assess specific risks. If the purpose of a research call is understood to

be risk assessment, then it should be carried out in accordance with general principles and best practice of risk assessment, which should include as a priority the robust consideration of uncertainty assessments.

As mentioned above, this is only proposed as an approach for some research calls, not for all of them. And I agree, it is not a panacea.

5 **Change 2b**

New text, p7 lines 11-16: And for some thresholds, the probability of crossing them is not be quantifiable. In these cases, the best approach may be to use qualitative descriptors of likelihood – such as the ‘very unlikely’, ‘possible’, and ‘likely’ used in the IPCC Working Group I’s assessment of the likelihood of specific abrupt and irreversible changes in the climate system occurring during the twenty-first century (IPCC, 2013), or the ‘low’, ‘moderate’, ‘substantial’, ‘severe’, and ‘critical’ used by
10 the UK government in its assessments of the likelihood of a terrorist attack (UK Government, 2019) – incorporating these into an assessment of how the risk will change as a function of time.

Comment 3a

This article points to an important improvement that can be made to the representation of risks related to climate change to
15 ensure that risk assessment reports speak more clearly and directly to the concerns of policy makers. Below I will describe some additional psychological factors in support of the boiling frog’s case, as well as give pointers for its successful implementation.

An empirical imperative for testing alternative climate risk representations

Studies in the psychology of risk have shown that its perception can be dramatically affected by how a particular risk is
20 presented or formulated (Slovic, 2010). Most formal risk assessments are highly cognitive endeavours. They rely on complicated mathematical analysis and are presented using formalisms that require expert knowledge to be understood. Beyond this aspect of risk known as risk-as-analysis, there are two related psychological dimensions to risk perception. Firstly, the risk-as-feeling hypothesis states that some risk problems generate an affective response while others might not. This affective response is sometimes an overt feeling of fear or anxiety, but often a more subdued “background” feeling
25 (Loewenstein et al. 2001). Secondly, it appears that some cognitive representations of risk problems are intuitively evaluable while others are not (Cosmides Tooby, 2008). This means that they can be evaluated using fast, automatic cognitive processes, rather than requiring slower, deliberative reasoning. Both risk-as-feeling and intuitive evaluability are crucial in guiding people’s day-to-day decision making. By extension, these psychological factors also affect how particular problems “jump
out” at politicians and policy makers, and thus influence how they trade-off giving attention to some risks over others.

30 Whereas some risk representations may be naturally “intuitively evaluable” by many people, other representations only become so through prolonged exposure and expertise in a particular domain. The “Hazards - Vulnerability - Exposure” risk framework of the IPCC WG2 Fifth Assessment Report is a likely case in point (IPCC, 2014). Journalism research following the release of the report has shown that its risk framing was not picked up by the media - with the notable exception of some business media (Painter, 2015). Although the primary audience of the report is policy makers, the fact that the risk language used in the

report's press release did not make it into articles about it suggests that it did not speak to the intuitions of the journalists involved. This can be seen as a proxy for other audiences outside of scientists and policy makers closely involved in the production of the WG2 report. These 'other audiences' may include policy makers in infrastructure, transport or treasury roles who will in future be affected by climate change risks, but whose professional background may lie in other domains such as economics or engineering. To make climate change risk representations gain widespread traction outside of the narrow band of scientists and climate policy makers, they will need to be made more intuitively evaluable to those other audiences.

How to make climate risk representations intuitively evaluable to policy makers?

The degree to which a complex problem is intuitively evaluable is related to a person's "lived expertise" in that problem domain. Policy makers often have expertise in one or more specific domains, in which they may have a limited number of concrete decision concerns. These concerns can be called the 'risk currency' of the policy maker, i.e., the measures or quantities that fall within their remit to keep below certain levels or between certain boundaries. The author's story (in AC1) of "faeces floating the in the street" is one such case. In this particular example, the scientists involved found it straightforward to produce a graph in the risk currency of the policy makers, which was hence intuitively evaluable to them, given their expertise.

As both reviewers rightfully point out, defining non-arbitrary thresholds may not be feasible or appropriate for all climate science areas or for all policy concerns, nor may the probability of surpassing a non-arbitrary threshold be calculable. To give an example of the latter: when a group of scientists studying Atlantic Ocean currents recently briefed Members of the European Parliament about their research, the most pressing question the MEPs wanted answered was what the research meant for the flow of refugees across the Mediterranean! Here, the risk currency of the politicians is one that the science community may never be able to address directly.

20 **Response 3a**

I'm grateful for these comments which provide additional insight into the issues raised in the paper.

I think the distinction between the information content of a risk assessment ('risk-as-analysis') and the presentation of a risk assessment ('risk-as-feeling' and 'intuitive evaluability') is a helpful one to make.

The main argument of my paper is that within the relatively narrow scope of the information content of the risk assessment ('risk-as-analysis'), there is room for improvement. If the variables of probability, impact and time have not been explored enough to bring the largest risks to light, then the risk assessment has omitted information that is likely to be of great relevance to the decision-maker. I propose identifying a non-arbitrary threshold of impact, and then assessing its probability over time, mainly because this seems likely to ensure that the largest risks are considered – and are not left hanging somewhere beyond the end of the x-axis or in the invisible margins of a shaded band of uncertainty. I believe that on this basis alone, there are advantages to this approach.

At the same time, I think it is quite possible that the probability over time approach also has an advantage in intuitive evaluability, and if so, this would be another argument in its favour. The discussion comment points out that 'The degree to which a complex problem is intuitively evaluable is related to a person's "lived expertise" in that problem domain'. This could explain the value of the 'experiential' threshold, as used in Christidis, Jones and Stott, 2015: once an extreme weather event

has been experienced, it becomes part of the ‘lived expertise’ of all those it affected. Extreme event attribution studies use this effect to give greater social salience to the reporting of climate change in the present. Risk assessments using experiential thresholds of impact could do the same for the communication of possible climate changes of the future.

Change 3a

5 [No change made in response to this comment.]

Comment 3b

What this example and the related concerns expressed by one of the reviewers (see comment RC3) make clear is the need for appropriate “co-production” of research questions between policy makers, scientists and research funders (De Meyer et al., 10 2018). Even if scientist may not be able to quantify the risk of refugees directly, they might find - in conversation with policy makers - intermediate ways to explain the possible multiplier effect of climate change on their concerns. To do this successfully for all policy questions, it will simply not be enough for policy makers to tell funders what questions they want answered; nor can funders define in isolation what research is required; nor can scientists define what thresholds are relevant to the risk 15 currencies of different policy makers without having meaningful conversations with them. The responsibilities for fixing this potential mismatch of knowledge production and knowledge requirements are fragmented, and will need joined-up and sustained efforts to resolve (ibid.). A first step to a solution is the recognition of the problem, and the author’s contribution to this could not have come at a better time.

Response 3b

Finally: I agree with the point about the need for co-production of climate change risk assessments. The identification of a 20 meaningful non-arbitrary threshold of impact is often likely to be possible only through dialogue between scientists and decision-makers. It is useful to recognise that this dialogue does not necessarily happen by itself. Consequently, dedicated and deliberate processes may be needed to produce information that is appropriate for the purpose of risk assessment.

Change 3b

New text, p7 lines 20-23: Overcoming this obstacle is unlikely always to be as simple as asking policymakers what it is that 25 they are most worried about. Without enough information to begin with, how can they know? An iterative process of ‘co-production’ of the risk assessment may be ideal, but the responsibility for coordinating such a process is not clearly owned by any one party (De Meyer et al., 2018).

References

- 30 [1] Lowe, J.A. and Bernie, D., ‘The impact of Earth system feedbacks on carbon budgets and climate response’, *Philosophical Transactions of the Royal Society*, Vol 376 Issue 2119 (2018) <https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0263>
- [2] Hare, D., ‘An Actuarial Perspective’ in King, Schrag, Zhou, Qi and Ghosh, ‘Climate Change: A Risk Assessment’ (2015) pp20-21 <http://www.csap.cam.ac.uk/links/13/1032/>

- [3] Table 12.4 on p1115 of Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. 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 [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_Chapter12_FINAL.pdf
- [4] UK Government, 'Terrorism and national emergencies' <https://www.gov.uk/terrorism-national-emergency>

10 **Telling the boiling frog what he needs to know: why climate change risks should be plotted as probability over time**

[version with changes highlighted]

Simon Sharpe¹

¹Institute for Innovation and Public Purpose, University College London, Gower Street, London, WC1E 6BT, United Kingdom

15 *Correspondence to:* Simon Sharpe (simonsharpe79@hotmail.com)

Abstract. Humanity's situation with respect to climate change is sometimes compared to that of a frog in a slowly-boiling pot of water, meaning that change will happen too gradually for us to appreciate the likelihood of catastrophe and act before it is too late. I argue that the scientific community is not yet telling the boiling frog what he needs to know. I use a review of the figures included in two reports of the Inter-governmental Panel on Climate Change to show that much of the climate science communicated to policymakers is presented in the form of projections of what is most likely to occur, as a function of time (equivalent to: 'in 5 minutes' time, the water you are sitting in will be two degrees warmer'). I argue from first principles that a more appropriate means of assessing and communicating the risks of climate change would be to produce assessments of the likelihood of crossing non-arbitrary thresholds of impact, as a function of time (equivalent to: 'the probability of you being boiled to death will be 1% in five minutes' time, rising to 100% in twenty minutes' time if you don't jump out of the pot'). This would be consistent with approaches to risk assessment in fields such as insurance, engineering, and health and safety. Importantly, it would ensure decision-makers were informed of the biggest risks, and hence of the strongest reasons to act. I suggest ways in which the science community could contribute to promoting this approach, taking into account its inherent need for cross-disciplinary research and for engagement with decision-makers before the research is conducted, instead of afterwards.

20

25

1 Introduction and argument from first principles

As the conceptual framework of ‘risk assessment’ is increasingly applied to climate change, the need to consider low probability, high impact risks (‘tail risks’) is often pointed out (Weitzman, 2011; IPCC, 2014a). What is not so often mentioned is that this principle is a subsidiary of a more general principle, which is perhaps taken to be self-evident: that a risk assessment should consider the biggest risks. In the case of a climate change risk assessment, how should we ensure that it does so?

If the magnitude of a risk is a function of probability and impact, then a risk assessment must consider three fundamental variables: probability, impact and time. To be sure of identifying the biggest risks, all three variables must be explored fully. But to fully explore any two of them, the third must be held constant. So the question is which choice of constant will lead to the fullest assessment of the risks.

If a risk is unchanging over time (at least to a rough approximation), then the answer is simple: hold time as constant by fixing a duration of interest, and then plot impact against probability. An earthquake risk graph as shown in Fig. 1 is such an example. It shows the full range of probabilities and impacts, from which the biggest risks can be understood. The time period is arbitrary, but changing it would not provide any significant further information.

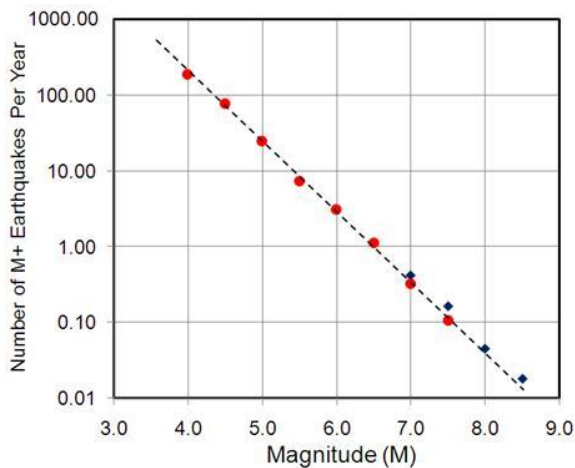
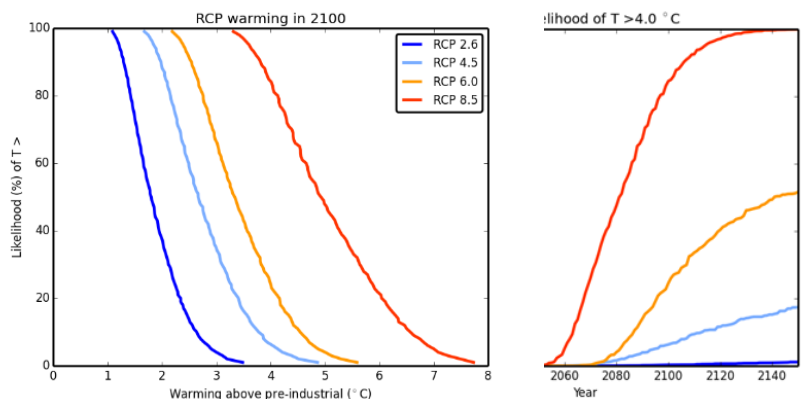


Figure 1: Example of probability / impact graph for a risk unchanging over time: plot of frequency / magnitude of earthquakes in the Chile area (Braile, 2010)

For risks that change over time, the choice is not so obvious. If time is held constant at a fixed point, then the full range of probabilities and impacts can be explored at that point, but bigger risks that may occur at different times will not be visible. If probability is held

constant and impact plotted against time, then bigger risks may be omitted either because they correspond to a probability other than that which has been chosen, or because they would occur at a later time than is shown on the x axis. (Shaded bands illustrating uncertainty in impact can bring a broader range of risks into view, but still provide no



guarantee that the biggest risks will be visible.) Similarly, if impact is held constant and probability plotted against time, then bigger risks may be omitted if they have larger impacts or occur at later times. These differences are illustrated in the three different plots of global temperature increase, probability and time shown in Fig. 2.

Figure 2: Three configurations of probability, impact and time, where temperature increase = ‘impact’: a) impact / time, (fixed probability); b) probability / impact, (fixed time); c) probability / time, (fixed impact); (Lowe and Bernie, 2015)

The difference between these three approaches lies in the relative arbitrariness of their fixed points. Fixing time makes little sense because while there is an obvious starting point (the present), there is no obvious end-point or discontinuity. Any fixed future point in time (e.g. the year 2100) is arbitrary. Probability has some interesting values, such as 0.5: the point at which something becomes more likely than not. But it is not clear that any particular value has special relevance for a risk assessment: the biggest risks could occur at 1%, 5%, or 63%. Furthermore, the fact that the range of probability is bounded at both ends – by 0 and 1 – makes it particularly well-suited to being one of our axes.

Impact, by contrast, may well have some fixed points that are not arbitrary, but highly meaningful. This can be seen in examples of regulations for the structural integrity of buildings in earthquakes, the capital reserve requirements for insurance firms, and the health and safety standards for people at work, which set maximum tolerable probabilities for building collapse, insurance firm insolvency, and worker death respectively. In each of these cases, the chosen probability is arbitrary, but the chosen fixed point of impact is not. For the building, insurance firm, or worker, the impacts chosen represent ‘worst case’ outcomes, beyond which no greater impact would be possible. On the range of possible severities of impact, these points represent discontinuities. Where such discontinuities can be identified, it may be most useful for a risk assessment to plot the probability of encountering them as a function of time.

To illustrate the relevance of this for risk assessment, consider the proverbial frog in a slow-boiling pot of water. If the frog asks his science adviser for advice, and is told that in five minutes, the water will be warmer by 2°C plus or minus a degree or two (illustrated with an impact over time graph), he may decide there is no compelling reason for him to get out. If instead he asks first what is the worst that could happen, and then how likely this is, his adviser will tell him that he could be boiled to death, and that while the probability of this is low within the next five minutes, it is rising over time, and at some point it will become more likely than not. Presented with the graph of probability of boiling as a function of time, the policy conclusion for the frog will be relatively clear.

Climate change has no single, obvious, ‘boiling frog’ scenario. There is no temperature threshold within which we are safe, and beyond which we are all cooked. Still, there is no reason why a similar approach could not be taken to assess a range of climate change risks. For example, Sherwood and Huber (2010) estimated that climatic conditions exceeding human physiological tolerance for heat stress would occur in parts of the world when temperatures rose 7°C above the late-20th century

average.¹ This may be compared, albeit not exactly, with the probability of exceeding 7°C above pre-industrial, estimated as a function of time for RCP8.5 by Jason Lowe and Dan Bernie, using a simple climate model to represent the climate sensitivity probability distribution function of the CMIP5 model ensemble (Fig. 3).

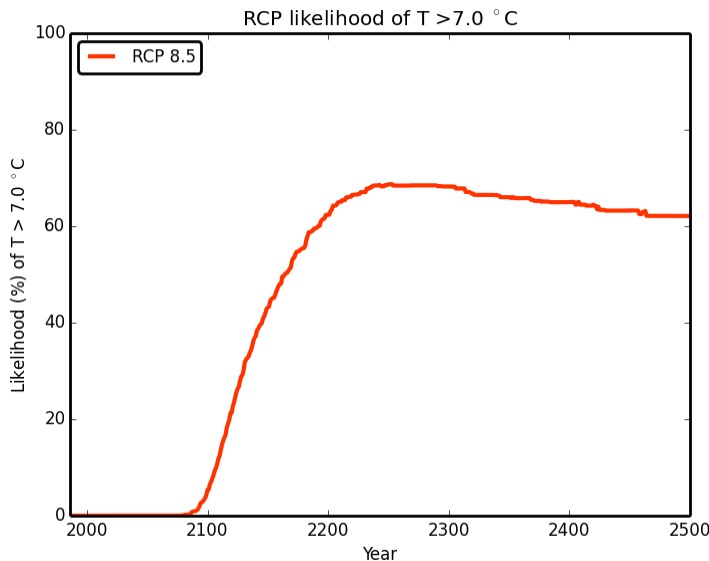


Figure 3: A ‘boiling frog’ example: the probability of global mean temperature exceeding 7°C above pre-industrial, as a function of time, for RCP8.5 (Lowe and Bernie, 2015)

While the probability of exceeding this threshold of temperature rise is low within the arbitrarily-defined time period of this century, it appears to rise rapidly thereafter, until it becomes more likely than not. Shown this way, a risk that might be assumed in the short-term to be negligible is seen in quite a different light.

2 Review of figures in IPCC reports: a predominance of projections of most likely impacts

Despite the argument set out above, it appears that the majority of graphs of future climate change impacts take the form of impact over time. By a rough count, the Working Group II contribution to the IPCC’s Fifth Assessment Report (IPCC, 2014b; IPCC, 2014c) contains some 26 figures featuring graphs of impact over time (including eight where a time-dependent variable such as temperature increase or emissions may be considered a proxy for time on the x axis). It contains a similar number of figures featuring maps, which when presented individually show impacts at a fixed point in time, and when presented in time-series are equivalent to graphs of impact over time. The report has no figures containing graphs of probability over impact. It has only four figures showing graphs of impacts as probability over time (with proportion or frequency taken as proxies for probability), and two map sequences that can be interpreted in a similar way. Only two of these probability over time graphs, and one map sequence, clearly relate to relatively non-arbitrary thresholds of impact – defined in terms of their physical effect rather than in relation to their historical likelihood. The map sequence shows how the proportion of days in the year with temperatures above 40°C – when severe heat wave consequences are experienced – could increase over time in Australia. The

¹ Analysis by Dr. Tord Kjellstrom, Professor Alistair Woodward, Dr. Laila Gohar, Professor Jason Lowe, Dr. Bruno Lemke, Lauren Lines, Dr. David Briggs, Dr. Chris Freyberg, Dr. Matthias Otto, and Dr. Olivia Hyatt in King et al., 2015 (pp57-63) showed that in some regions this threshold may begin to be crossed at significantly lower values of global average temperature increase.

two figures with graphs both relate to the risks to corals. One of these is reproduced here as Fig.4; it shows how the proportion of coral grids with degree heating months above threshold values for mass bleaching and mortality could change over time.

5

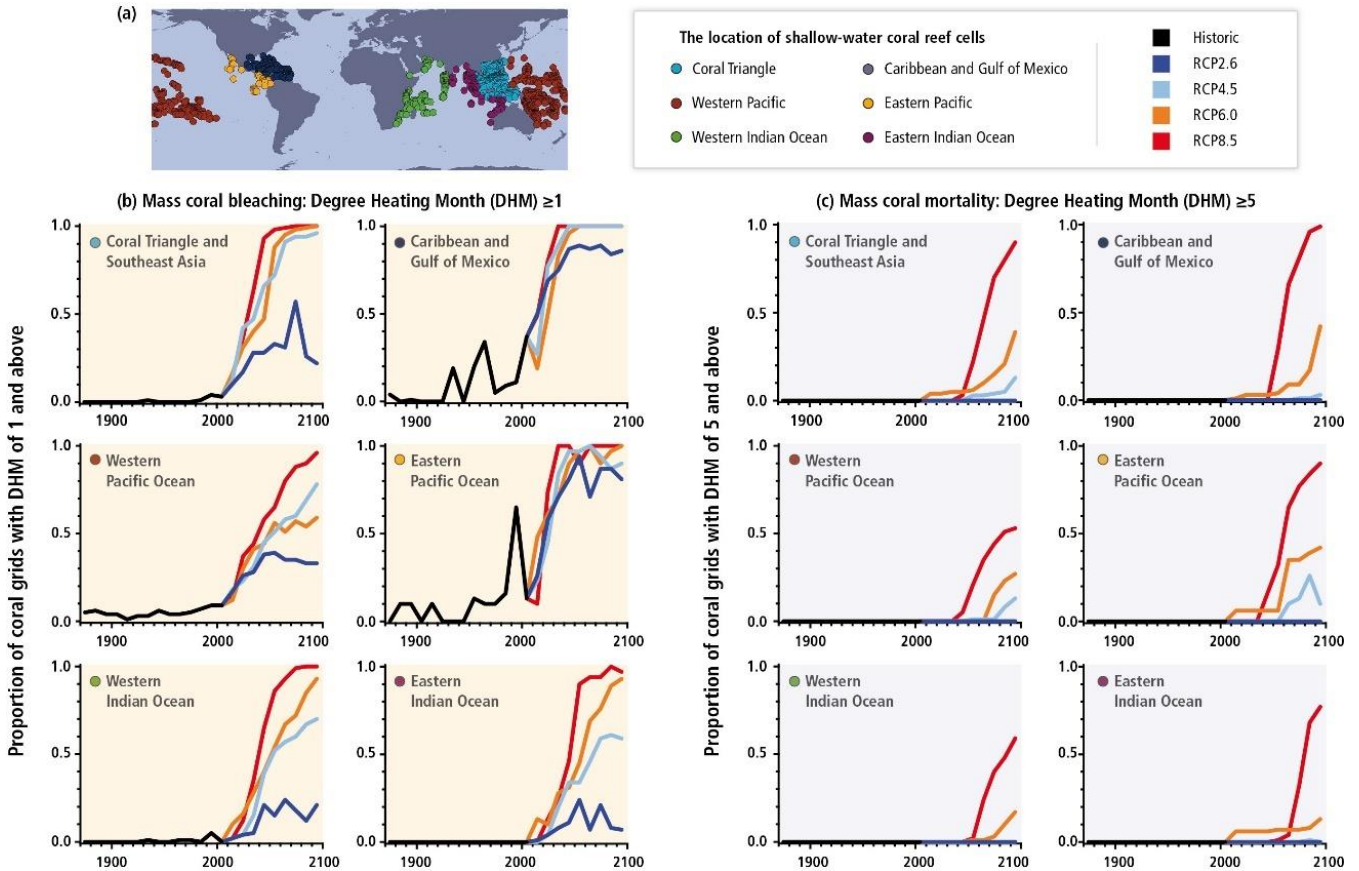


Figure 4: The risk of mass coral bleaching / mortality presented in the form of probability / time² (Fig. 30-10 in IPCC, 2014d).

² Original caption: Annual maximum proportions of reef pixels with Degree Heating Months (DHM, Donner et al., 2007) for each of the six coral regions (a, Figure 30-4b)—(b) DHM ≥ 1 (used for projecting the incidence of coral bleaching; Strong et al., 1997, 2011) and (c) DHM ≥ 5 (associated with bleaching followed by significant mortality; Eakin et al., 2010)—for the period 1870–2009 using the Hadley Centre Interpolated sea surface temperature 1.1 (HadISST1.1) dataset. The black line on each graph is the maximum annual area value for each decade over the period 1870–2009. This value is continued through 2010–2099 using Coupled Model Intercomparison Project Phase 5 (CMIP5) data and splits into the four Representative Concentration Pathways (RCP2.6, 4.5, 6.0, and 8.5). DHM were produced for each of the four RCPs using the ensembles of CMIP models. From these global maps of DHM, the

Figure 4, the exception that proves the rule, does a good job of communicating the risk: it makes it quite clear that on a high emissions pathway, it will be only a matter of time until most of the world's corals are extinguished.

- 5 A review of the IPCC's Special Report on Global Warming of 1.5°C (IPCC, 2018) suggests that this pattern has changed little over the intervening four years. Impact over time graphs and maps still predominate, although temperature (a time-dependent variable) is used for the x axis rather than time itself – perhaps reflecting the report's purpose of demonstrating the differences between warming of 1.5°C and 2°C. While the proportion of probability over time figures has increased – to four out of seventeen relevant figures – only one of these clearly relates to a non-arbitrary physical threshold of impact. (This shows the
- 10 fraction of global natural vegetation at risk of severe ecosystem change as a function of global mean temperature change – Fig. 3.16 in IPCC, 2018).

3 Discussion: the opportunity and need for more assessments of probability over time

There are many ways that the stock of probability over time assessments could be expanded. Non-arbitrary fixed points of impact can be defined in relation to several different kinds of thresholds:

- 15
- Physical: the height of sea level that puts an island under water;
 - Biophysical: the degree of heat and humidity that exceeds human physiological tolerance*; or the temperature that exceeds a crop's tolerance*;
 - Biochemical: the degree of acidity that prevents a shellfish from forming a shell;
 - Socioeconomic: the quantity of per capita water resources required to meet basic human needs*; the daylight hours below dangerous levels of heat stress required for a subsistence agriculture lifestyle to remain viable; or the height

20 of sea level at which it becomes less costly to relocate a coastal city than to continue to protect it against flooding*;
 - Experiential: the impact of a past event whose damage is well understood, e.g. a storm surge equal to that of Superstorm Sandy, or a European heat wave equal to that of 2003 (as illustrated in Christidis, Jones and Stott, 2015);
- 25
- Political: an agreed value, such as the 2°C target warming limit*.

Clearly, these different kinds of threshold vary in their objectivity, and in other ways likely to have implications for risk assessment. Socioeconomic thresholds may be more possible to overcome through adaptation than those defined by physical or biochemical properties alone. Experiential and political thresholds may have only transient value – useful for as long as the

annual percentage of grid cells with DHM ≥ 1 and DHM ≥ 5 were calculated for each coral region. These data were then grouped into decades from which the maximum annual proportions were derived. The plotted lines for 2010–2099 are the average of these maximum proportion values for each RCP. Monthly sea surface temperature anomalies were derived using a 1985–2000 maximum monthly mean climatology derived in the calculations for Figure 30-4. This was done separately for HadISST1.1, the CMIP5 models, and each of the four RCPs, at each grid cell for every region. DHMs were then derived by adding up the monthly anomalies using a 4-month rolling sum. Figure SM30-3 presents past and future sea temperatures for the six major coral reef provinces under historic, un-forced, RCP4.5 and RCP8.5 scenarios.

past event is remembered, or for as long as the policy holds. But what all these thresholds have in common is that they are relevant to policymakers, because they are defined in terms of what it is that we collectively wish to avoid. Even a subjectively-defined threshold of impact, chosen for its social relevance, is less arbitrary in this sense than a fixed point of probability or of time.

5

The report ‘Climate Change: A Risk Assessment’ (King et al., 2015) demonstrated the feasibility of the probability over time approach by presenting some illustrative studies and discussion of the probabilities of exceeding a range of thresholds, including those marked with an asterisk in the list above, as a function of time (or time-dependent variables) for selected locations and scenarios. These examples showed clearly that on certain pathways, the things we wish to avoid may become highly likely. The response to the report was encouraging: at the Chatham House climate change conference of 2015, attended by members of the international climate policy community, a presentation of the report’s findings was voted the second-most valuable presentation of the conference, from a field of some thirty expert speakers.

10

The probability over time approach is not without its challenges. Most thresholds do not have a single ‘correct’ definition. For example, a temperature tolerance threshold may be biophysical, but how long above that temperature should be considered ‘too long’ is a matter of expert judgment. Many thresholds are specific to their location, meaning that a national or global risk assessment performed this way may need to be made up of a diverse range of quite distinct studies, rather than using the kind of consistent data that allows for aggregation. And for some thresholds, the probability of crossing them is not be quantifiable. In these cases, the best approach may be to use qualitative descriptors of likelihood – such as the ‘very unlikely’, ‘possible’, and ‘likely’ used in the IPCC Working Group I’s assessment of the likelihood of specific abrupt and irreversible changes in the climate system occurring during the twenty-first century (IPCC, 2013), or the ‘low’, ‘moderate’, ‘substantial’, ‘severe’, and ‘critical’ used by the UK government in its assessments of the likelihood of a terrorist attack (UK Government, 2019) – incorporating these into an assessment of how the risk will change as a function of time.

15

20

25

The experience of producing the abovementioned report suggested that the most significant obstacle to adopting the probability over time approach was not any difficulty with the science, but the need to start – *before doing any science* – with a subjective question: ‘what is it that we wish to avoid?’ Overcoming this obstacle is unlikely always to be as simple as asking policymakers what it is that they are most worried about. Without enough information to begin with, how can they know? An iterative process of ‘co-production’ of the risk assessment may be ideal, but the responsibility for coordinating such a process is not clearly owned by any one party (De Meyer et al., 2018).

30

It is therefore worth considering how each part of the climate science community can contribute to bringing these assessments into being. Contributors to the IPCC’s Working Group II could conduct risk assessments using impact thresholds of the kinds described above. Some of these are likely to highlight thresholds of temperature or sea level rise at which non-linear increases

in risk take place, and these thresholds could in turn be assessed in the form of probability over time by contributors to IPCC Working Group I. Working Group I might also be able to assess the risks of ‘large-scale singular events’ in this way. National climate change risk assessments could identify thresholds of impact of particular relevance to the country concerned. Researchers conducting extreme event attribution studies could run their models forward in time – to show how the probability of crossing each newly-formed experiential threshold will continue to increase in the future. **Research funders could play an influential role by structuring research calls in ways that require co-production with decision-makers, inter-disciplinary collaboration, and the application of general principles of risk assessment.**

4 Conclusion

10 The risks of climate change can be understood more clearly when research starts by identifying what it is that we most wish to avoid, and then assesses its likelihood as a function of time. By providing a clearer picture of the overall scale of the risks of climate change, such assessments could help inform the most important decision of all: how much effort to put into reducing emissions.

15 **The proposal is certainly not that all climate science should be done in this way. Fundamental research is indispensable; and there are many ways of communicating risks. The suggestion is that more research could be done expressly for the purpose of risk assessment than is done at present, and that a deliberate approach should be taken to identifying and assessing the biggest risks.** Decarbonising the world economy will not be as easy as jumping out of a pot. That makes it all the more important that no opportunity is missed to communicate the severity of the risks to those in charge. The water is already
20 getting warm.

Data availability

The data underlying Section 2 of this paper is stored with open access at the National Geoscience Data Centre (item 125176), and is available for download at <https://www.bgs.ac.uk/services/ngdc/accessions/index.html#item125176>

Competing interests

25 The author declares that he has no conflict of interest.

Acknowledgements

Especial thanks to Jason Lowe and Alistair Woodward, for sharing their knowledge and advice so generously, without which this work would not have been possible. Sincere thanks also to Sir David King, Chris Rapley, Kris de Meyer and Rowan Sutton, for their advice and support. Thanks also to the reviewers for their constructive and helpful comments.

5 References

- Braile, L., Feb. 27, 2010 M8.8 Chile Earthquake: Seismic/Eruption Views, Seismicity, Main Shock-Aftershock Sequence, Purdue University, (2010), available at: <http://web.ics.purdue.edu/~braile/edumod/chile/chile.htm> , 2018.
- Christidis, N., Jones, G.S., and Stott, P.A., Dramatically increasing chance of extremely hot summers since the 2003 European heatwave, *Nature Climate Change* 5, 46–50, doi:10.1038/nclimate2468, 2015.
- De Meyer, K., Howarth, C., Jackson, A., Osborn, D., Rose, L., Rapley, C., Sharpe, S., and Welch, K., Developing Better Climate Mitigation Policies: Challenging current climate change risk assessment approaches, UCL Policy Commission on Communicating Climate Science, Report 2018-01, available at: https://www.ucl.ac.uk/public-policy/sites/public-policy/files/risk_workshop_report_final_0.pdf , 2018.
- IPCC, 2013: Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. 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* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1115pp., 2013.
- IPCC, 2014a: Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.), Summary for policymakers. 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*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 9, 2014.
- IPCC, 2014b: Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.), *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*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1132 pp., 2014.
- IPCC, 2014c: Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.), *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 688 pp., 2014.
- IPCC, 2014d: Hoegh-Guldberg, O., R. Cai, E.S. Poloczanska, P.G. Brewer, S. Sundby, K. Hilmi, V.J. Fabry, and S. Jung, The Ocean. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of*

Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1655-1731, 2014.

5

IPCC, 2018: O. Hoegh-Guldberg, D. Jacob, M. Taylor, M. Bindi, S. Brown, I. Camilloni, A. Diedhiou, R. Djalante, K. Ebi, F. Engelbrecht, J. Guiot, Y. Hijikata, S. Mehrotra, A. Payne, S. I. Seneviratne, A. Thomas, R. Warren, G. Zhou, 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, sustainable development, and efforts to eradicate poverty [V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, T. Waterfield (eds.)]. In Press. 2018.

10

15 King, D., Schrag, D., Zhou, D., Qi, Y., and Ghosh, A., Climate Change: A Risk Assessment, Cambridge University Centre for Science and Policy, available at: <http://www.csap.cam.ac.uk/media/uploads/files/1/climate-change--a-risk-assessment-v11.pdf> , 2015.

20

Lowe, J. and Bernie, D., Met Office Hadley Centre, AVOID Programme, Global Temperature Increase, in: King, D., Schrag, D., Zhou, D., Qi, Y., and Ghosh, A., Climate Change: A Risk Assessment, Cambridge University Centre for Science and Policy, 50-55, available at: <http://www.csap.cam.ac.uk/media/uploads/files/1/climate-change--a-risk-assessment-v11.pdf> , 2015.

25

Sherwood, S.C. and Huber, M., An adaptability limit to climate change due to heat stress, PNAS, 107 (21), 9552-9555, doi:10.1073/pnas.0913352107, 2010.

UK Government, Terrorism and national emergencies, available at: <https://www.gov.uk/terrorism-national-emergency> , 2019.

30 Weitzman, M.L.: Fat-Tailed Uncertainty in the Economics of Catastrophic Climate Change, Review of Environmental and Economic Policy, 5 (2), 275-292, doi:10.1093/reep/rer006, 2011.