Re-review of "Ice Melt, Sea Level Rise and Superstorms: Evidence from Paleoclimate Data, Climate Modelling and Modern Observations that 2C Global Warming is Highly Dangerous" By James Hansen and colleagues

Peter Thorne, Maynooth University
06/12/2015

As noted in my original review, the high level of publicity and large volume of comments has made this paper highly atypical and likely problematic to review. My opinion remains that this is the case. Having read the majority of the comments and responses, in addition to the revised manuscript, I shall undertake the re-review also in an atypical fashion by:

1. Providing opening remarks to highlight key thematic areas of remaining concern which relate to the process, the paper structure and presentation, and the science;
2. Highlighting a number of additional specific aspects of the various initial reviews that I feel have either been inadequately responded to, or not sufficiently incorporated in the revisions to date;
3. Performing a more traditional review of the submitted redrafted manuscript content for remaining issues, in the expectation that the authors and Editor may in their collective judgement wish to proceed further.

For ease of reading I separate each distinct segment by a page break.

Concerns about 'gate-keeping' are foremost in my mind when reviewing papers. Certainly, efforts should stringently be made to ensure that the literature is reflective of the entire range of scientifically valid opinions backed up by appropriate experimental findings. In that context, peer-review is a necessary, but not sufficient, step to eventual broad acceptance of a piece by the scientific community. Therefore, arguably, the default is eventual acceptance of any reasonable and reasoned submission that is backed up by scientifically valid evidence.

Equally, peer review exists to ensure some minimum standard scientifically is upheld; that the literature reflects only scientifically plausible propositions, and that any published paper is written in a style commensurate with the journal and the field (nomenclature, caveats, discussion etc.).

I will generally recommend publication of pieces with which I disagree so long as a sufficient level of scientific justification is given that convinces me that the findings are at the very least not impossible, even if I personally believe them implausible, and the text is couched in an appropriate manner. This is particularly so when the paper does not entirely sit within my areas of expertise.

That all being said, at this time I remain in a position whereby I am unable to provide in good faith to the Editor a firm recommendation upon whether to publish or not owing to a number of, in my view, outstanding issues. I would need to see a revised version along with a point-by-point response to my review herein (including all three sections outlined above), to be able to provide such a recommendation.

I would bring to the Editor's attention that the authors have clearly undertaken a considerable effort to redraft the paper. The new version is certainly a demonstrable improvement in terms of readability, and is duly recognised for being so. It has also taken into account a subset of the formal reviews and short comments received.

Opening remarks

Tenor of ACPD responses

I find the authors' choice of style of responses to the short comments and solicited reviews posted at ACPD disappointing in very many cases. As I noted in my review; there were many that were of limited utility even by then (and more followed) and to which short, but polite, responses would suffice. However, there were of the order 10 germane inputs by experts. Many of these were responded to in what I can only describe as an unprofessional manner. A manner that, sadly, all too easily could be implied to verge on disrespect both for the authors of the short comments, and for the ACP journal and its review process.

These contributors, without fail, went to the considerable effort to read the ACPD manuscript and put forward substantive comments for due consideration by the authors, in the reasonable expectation that they would be taken seriously and help the authors improve the paper. Typically this would have taken several hours of entirely voluntary effort on their part. It is hard enough for journal editors to muster reviews, even when solicited. It is rarer still for EGU journals to get multiple perceptive and valuable unsolicited comments as was the case here, admittedly amongst a number of less germane comments. When the authors are seen to not treat that seriously, the whole system is put at potential risk.

It is, in my view, beholden upon an author team of such seniority within the community, on a paper with such visibility, to set an exemplar of how to do this. In my judgement the current responses, including the public record response to reviewers, do not attain the required standard. They are, with one or two notable exceptions, editorialising and ignore the substantive points raised. There are no point-by-point responses to all points raised that enable the reviewers (both solicited and unsolicited) to understand what changes were made in response to each comment and why (or why not), as would be typical for responding to reviewers, either in OA or a more traditional review process.

Were there the potential still available at this stage to supplement existing responses with more polite, comprehensive and germane responses, I would in the strongest possible terms urge the journal to take the extraordinary step of reopening to the author response step and enabling this. Were this step re-opened I would then urge the author team to write more appropriate responses that supersede the current response of record, and respond to each point raised by reviewers in a substantive and constructive manner. The authors should in such a case be given the time needed to respond in a substantive way to all the points raised in each meaningful review.

If such re-opening to the responses phase is not allowed, then new responses should be appended to this response for the following comments:

- Collaborative comment led by S. Drijfhout
- Collaborative comment led by G. Flato
- Collaborative comment led by W. F. Ruddiman
- Comment by M. de Rougement

- Comment by M. Wehner
- The first comment by D. Berner
- The two initial formal reviews

In all cases the original response was either substantively incomplete or of an inappropriate tone (or both). I would note that in the interests of time my own assessment, as part of this re-review, of the initial reviews and responses was necessarily incomplete. So, I would welcome the editor suggesting any additional short comments to which they feel an updated response would be warranted.

Were the paper to be accepted, my understanding is that the intermediate versions and closed period review communications become public. Therefore a more considered response to these various comments than is currently present would be on the public record either way that this is achieved. My preference would be for additional comments to be posted on the ACPD version, where people would naturally expect to find them.

The response to the collaborative comment led by M. Engel is a working example of the depth and tenor of response I would expect to see to the several substantive Short Comments and the two initial solicited reviews. Similar depth should be possible to all the substantive comments that were received, if the authors really wish their paper to be published.

I was particularly concerned at the response given to the entirely sensible review of S. Drijfhout and colleagues. In no sense is the following statement acceptable in the public record of review and response regarding a posited peer reviewed paper under **any** circumstances:

> *Hmmm, yes, I guess that we should not be worried about anything that happens 85 years from now- the dickens with those characters. The Dutch can migrate to Switzerland, after all.*

This is particularly so when it is made in the presumably full knowledge that several of the commenters are themselves Dutch (as, coincidentally, is the editor in charge of the paper).

The bottom line here is that this is not the school playground, and neither the commenters who provided substantive comments nor the authors are six years old. I expect this kind of thing of my kids. I do not expect this behaviour to be out there in the public domain for all to see amongst leading scientists in the field. It is unbecoming, unprofessional, and absolutely needs to be rectified. We, as a community, are better than this, and need to be seen as such.

To conclude this point, either new responses to the comments on the ACPD site (if permitted), or substantive responses as an appendix to this re-review response, are an absolute pre-condition personally to being able to recommend acceptance to the editor. Of course, the editor may take this under advisement, as is their right. But I feel very strongly that the current state of affairs leaves the journal, the authors, and the reviewers in an untenable position were the paper accepted without correcting this aspect of the public record.

Degree of editorialising / policy advocacy

This aspect remains an issue in my view from the initial draft per my first review, although it is somewhat better. Please note that I am not suggesting that scientists should obfuscate the policy implications of their work, and an assertion to that end in your public response to my review is unwarranted and unhelpful (see later).

There is a time and a place for discussing the policy implications of your work. My view remains strongly that an ACP paper is neither said time nor said place. The literature is where the science is discussed. Policy is far outside the described journal remit for ACP. If the authors wish to discuss policy they should submit to a journal that considers it within their topic domain rather than insist a journal extends its stated remit to suit their desire to discuss policy. I would therefore find it hard to recommend acceptance were the introduction to remain in given form, or for Sections 9 and 10.3 to remain at all. Both these latter sections should, in my view, be removed. I will now outline each of these three in turn with specific rationale and / or suggestions.

The introduction should be a more traditional-form introduction (an approach tried and tested for several Centuries across numerous scientific disciplines) that sets the scientific context for the study. Specifically it should outline:

1. The current state of the art in the science thematic areas to be covered in the paper. Here, for example, is the appropriate place to raise the SPM statements on SLR that are germane from AR5 as block quotes and without editorializing, or the mass balance closure condition common in CMIP5 model runs.
2. What is novel and new about this paper that means it constitutes a valuable scientific addition. In this, very briefly, the importance of exploring possible futures from multiple perspectives to *inform* policy makers in their decision making in a sentence or two could be touched upon without being policy prescriptive.
3. How the rest of the paper is structured, by enumerated section, and why, so that the reader knows what to expect and what the story you intend to tell is.

Much of this introductory material is instead in current Section 2, which should replace the current introduction and be expanded according to the above recommendations. All of current Section 1 I find to be too policy-orientated and, as such, I would be extremely averse to its retention in any form in an ACP publication.

Section 9 is a single page that is tangential, adds nothing scientifically to the paper, and involves a hefty amount of policy editorializing in Section 9.2. I would advocate its removal for overall paper readability and length regardless as it is entirely tangential to the main findings and conclusions. That it is also ringing journal scope alarm bells makes its removal, in my view, a pre-requisite for acceptance. It will help reduce the paper length and the general flow of the piece. There are many papers describing when the anthropocene may have begun etc. and there is no need, in my view, to include this subject, no matter how briefly, in

the current paper. Section 9 has no dependencies with other sections and is not highlighted in the conclusions. It can be deleted without any deleterious implications for the paper. I strongly suggest it be removed.

Section 10.3 is better than the text it replaces. However, it is still blurring the boundary between peer-reviewed literature and advocacy. Some of the section could remain if it removes the policy aspects but I would strongly prefer to see section 10.3 disappear. Were Section 10.3 to remain as is I would be extremely hard pressed to recommend acceptance, given the journal guidance to reviewers and stated journal scope.

There were a number of additional short passages that I shall return to in my more formal review at the end which in my view require modification.

Mischaracterisation of IPCC processes

It became obvious in the responses to the reviewers and comments that there are major misconceptions about IPCC being perpetuated. This is potentially particularly problematic given that the newly elected Chair of IPCC WG1 is amongst the authors. Almost all contentious points regarding IPCC are inappropriate regardless, but doubly so in the context of the role taken up recently by a co-author. IPCC does not, for example:

- Own, build, design or run climate models. So, IPCC models is a substantial misnomer, admittedly used unfortunately elsewhere on occasion. These are CMIP model runs, not IPCC model runs, and should be labelled as such.
- Undertake new scientific analyses. So there are no IPCC studies per se.
- Consist of a bunch of sheep who all think the same about all aspects of the science. So there is no collective noun that can pigeonhole the scientists as part of some collective that all think the same on all issues.

Yet in several places one, or sometimes more, of the above, or similar mis-conceptions, are promulgated by the authors both in the paper and in their review process communications on ACPD. All such cases in the manuscript need to be identified and rectified prior to being acceptable for publication. Most are highlighted in my traditional form review that follows at the end with specific suggestions as to how to modify.

Degree of certainty

Both in my review and a number of the Short Comments the point was raised (in many cases very strongly) that the results posited in the study were only a possible outcome. Furthermore, they likely arise from an extreme tail of the possible climates we shall experience in the 21st Century. As such, in the title, the abstract and the conclusions it is my view that the authors are still being highly unduly confident in their language. The title, abstract and conclusions are not sufficiently supported by the now somewhat more circumspect (appropriately so) text to remain as is. As noted in my original review extraordinary claims require extraordinary evidence. I do not see extraordinary evidence at play here

that supports a definitive assertion that the posited effects shall eventuate. I see at most indicative evidence, with substantial remaining uncertainties, that cannot rule out the eventuality posited. Given this, it is necessary to reframe the abstract, conclusions and title accordingly to be scientifically defensible and consistent with the underlying principal findings and caveats.

The title needs changing to reflect that the outcome is inherently uncertain if I am to be able to recommend acceptance. The easiest solution would be to insert Potentially before Highly Dangerous that would provide some sense of the uncertainty in the underlying analysis. More substantial changes would be along the lines of 'Exploring potential impacts of a 2C world using insights from paleoclimate records, modern observations and climate modelling' or 'Exploring the potential for tipping points in the climate system before 2C'. Basically, I think the title needs to reflect that the outcome is not deterministic and not guaranteed, even if we are foolish enough to stay on a carbon intensive pathway.

Both the abstract and the conclusions need to make clear that the evidence cannot rule out large-scale changes but that, equally, it is not a given that such changes shall occur. They need to better reflect that there remain substantial uncertainties and areas where further research is required to make definitive conclusions. Such revisions would be consistent with the underlying text and reflect the true state of scientific knowledge in the area.

Boulders

The storm-tossed boulders issue still is not satisfactorily resolved to my mind. This issue had been raised by numerous Short Comments, and in my first formal review. I have also discussed more latterly with recognised experts in extreme waves as part of this re-review. It is beyond question that these boulders are older than the substrate upon which they rest and must have been wave deposited, and I did not see any credible comments that suggested otherwise.

The critical unanswered questions therefore are: i) dating of deposition and ii) whether they were deposited by meteorologically driven waves and associated storm surges or a tsunami like set of waves, possibly associated with subsequent or temporally coincident orographic uplift if associated with a local-point source.

The dating uncertainty is hard to see how to resolve, but should be explicitly mentioned in the redraft. Specifically, the age estimates are based upon a single technique (amino acid racemization as far as I can ascertain), which has uncertainties, as do all dating techniques. The dating technique uncertainty presumably is broader than the 5e high sea-level stand period at a minimum. This implies we cannot be 100% sure these were deposited during 5e and the non-zero probability of this is not currently articulated in the manuscript. This dating uncertainty aspect should be clearly articulated if the section remains, and the potential implications for the authors' interpretation discussed. It is arguably not certain that these boulders were deposited specifically during stage 5e, yet this is the impression given to the readers.

On the wave source question, it must be possible to at least see whether we can rule out the meteorological storm induced hypothesis. And best scientific practice would require us to test against such a null before accepting it as a possible hypothesis. Despite suggesting such a test in my review no such analysis was undertaken in the revisions or alluded to in the responses to reviewers (myself or others).

The physics of the problem is fairly simple. Whatever tossed those relic boulders onto the present-day cliff top must have been powerful enough to 1. Dislodge them from the sea floor and 2. Toss them the at least 15-20m elevation gain locally to the datum at the time.

There are several ways of getting at the mass of these boulders. These are rocks that are approximated by a sphere of 12m at present day from the picture so have a volume of c.2000m^3 (4/3pi*r^3 where r =8m based on current size and assuming limestone being reduced by 125Kyr of chemical weathering reducing the size in the interim). Oolitic limestone when saturated has a density in the range 2.14-2.29Mg/m^3 (http://link.springer.com/article/10.1007%2FBF02595261) meaning that the saturated mass of these boulders would have been c. $4X10^6$Kg (this is 4 times the mass given on line 1168 of the revised manuscript). An alternative expert assessment given to me is that the largest boulder was $2.3x10^6$Kg. Regardless of the precise mass, this is very much larger than the boulders discussed in Cox et al. (2012), which were $40\text{-}80x10^3$Kg.

There is at least an order of magnitude difference to account for to suggest modern estimates on Aran, which gets some of the highest waves in present day climate from long-fetch high-powered mid-latitude N. Atlantic storm systems and where the cliff height is nearer 10 metres (I have been there myself), are a useful and useable analogue here. Meteorologically, fetch for Aran is likely substantially longer than for the location of the Boulders in question (I type this as the synoptic situation is a 4000Km straight run of SWrlys at 80-120Kph), arguably allowing more energetic ocean waves than meteorologically attainable in the sub-tropical location in question (wave energy being a function of wind speed, duration, and fetch), even with the strongest hurricanes. Furthermore, as far as I can tell, most of the discussion in Cox et al relates to movement of boulders in situ on the cliff top and not their deposition upon the cliff. Water mediated movement on a surface requires far less energy than transport and deposition onto the surface from a lower datum.

Coming back to the Bahaman site, the modern, apparently storm tossed, boulders are 1/10 the size or $1/1000^{th}$ the volume, and hence mass, of the relic boulders. Further, it is not clearly stated what the evidence is that these modern rocks are storm tossed (no reference is given). Even if they are storm tossed, it does not follow that the same processes can account for relic boulders three orders of magnitude heavier. Stating that they are $1/10^{th}$ the size is disingenuous without noting for the unwary reader that they are hence $1/1000^{th}$ the mass. At a minimum this needs to be acknowledged in a revised text.

1J is required to lift 1Kg 1m. To lift $4x10^6$Kg 15m would require $60x10^6$J or 60MJ (for a mass of $2.3x10^6$Kg c.35MJ, for a mass of $1x10^6$Kg 15MJ) of energy. The question then is whether it is plausible that meteorologically induced ocean waves could have had the requisite power to be able to dislodge, vertically transport, and then deposit such massive boulders. This is a question to which I do not have the requisite oceanographic knowledge to provide a definitive answer. My enquiries with relevant experts highlight that they would not absolutely rule it out without undertaking substantial calculations and fieldwork, but neither would it rule out a tsunami-based deposition mode.

Furthermore, the availability of such boulders as loose material, given the local environmental characteristics, without large-scale geologic disturbance as a mediator, is questionable. Such geologic upheaval may also have resulted in local orographic uplift reducing the work required to place the boulders atop the present-day cliffs in the first place. Overall, I therefore presently find the arguments for a local point-source tsunami deposition mechanism articulated by Engel et al. at least as plausible as the storm deposition posited by the authors.

The burden of proof lies with the authors here. Currently the evidence is being oversold in my opinion, that of a number of Short Comment submissions, and that of the relevant experts I reached out to. In my view, the authors need to at an absolute minimum provide a quantitatively based estimate of the waves and resulting storm characteristics that would have been required to deposit such large boulders rather than simply assert they were storm deposited. They also need to more clearly rule out a tsunami-based mode, or, if they can't, then to more clearly caveat that this alternative explanation is viable.

I would note that Michael Wehner's currently unanswered comment provides published evidence that we can expect hurricanes to increase in intensity with warming SSTs at most 10%. This provides a potential upper bound on storm intensity and size from which to infer meteorologically possible wave characteristics. Perhaps ancient storms were stronger still, but if so, the mechanisms would need elucidating with supporting references.

Without further quantitative analysis that proves the plausibility of a storm-based deposition vector or rules out better the tsunami deposition, I would have to advise the editor to accept only upon removal of this aspect from the paper, which unduly distracts and is to the opinion of myself, several commenters and experts I have reached out to, oversold. I find the chevrons evidence that follows more compelling regardless.

Paper structure

First, to be clear, the reorganisation of the paper has undoubtedly improved accessibility and readability. However, in reading the paper there are still issues over structure that serve to reduce its accessibility. I shall go through these in turn below.

1. The paper would benefit at the start of each major section from a brief paragraph outlining what the section shall discuss and how the section is structured (section x.y shall discuss z etc.). This text should replace the existing short paragraphs in many places at the end of existing sub-sections, that state what is to come in the next sub-section etc.. It would greatly enhance readability were this done.

2. There are many places where the text is describing aspects that should be in other sections for narrative continuity and readability. This is particularly prevalent in the modern observations section, where there are whole passages of text discussing exclusively either paleo data evidence or the models, without even a reference to the observations. This is not appropriate and reduces readability. Text should be reassessed throughout the paper and, if necessary, moved to the appropriate location. Such cases should ideally be reconciled with existing text to minimise paper length. The most egregious examples have been identified in my traditional long-form review at the end, but there were others.

3. It is unclear to me why the future projection modelling results are split into Section 3 and Section 4. It is also unclear to me what added value retaining the projection runs detailed in Section 3 brings to reader interpretation above and beyond the projection runs in section 4. It would significantly streamline the paper and aid reader interpretation of the modelling results if the authors considered and showed results only arising from Section 4 projection runs which presumably are the experiments they have most faith in, otherwise they would not have rerun the experiments.

   Can the sections be merged and results be shown for only the projection runs described in Section 4, supplemented by the remaining runs already described in Section 3? If not, then this is a reasonable question a reader would ask and the authors need to be very much clearer how the two sets of projection experiments are distinct and add value. Currently it appears entirely arbitrary.

   From my interpretation of the paper the sections should be merged and solely the runs from the experiments in section 4 discussed when considering 21st Century projections. Perhaps this is similar to the legacy that led to the split paleo sections in the original submission. While the evolution of the modelling study design may be of interest to the authors it is of limited utility to the reader – show the results you think are best and streamline the analysis accordingly. You can replot the 21st Century projections, using the runs described in Section 4, for all plots currently in Section 3 that relate to the model projection runs. My opinion is that this could save several pages of text and figures, and would serve to clarify the modelling section messaging for the reader substantively.

If the authors wish to retain the future projection model runs from both Section 3 and Section 4 then please come up with a way of helping the reader identify which model runs are being referred to when in subsequent sections. At present time it is impossible for the reader to understand which experiments with which forcings are being referred to in Sections 5 and 8 in particular. It would greatly simplify matters if only Section 4 future projection runs were included, and that is my firm recommendation to the Editor and authors at this time.

4. Similar to the above point – it is unclear to me what differentiates Section 7 from Section 6. In particular, Section 7 is a continuation of a discussion of the paleo evidence. Why the arbitrary split point here? It is not clear as it is written because Section 7 lacks a clear opening paragraph that states what it will consider, and why it is distinct. My reading is that Sections 6 and 7 are substantively similar and should be merged. In particular Section 7 appears to be largely a literature review / synthesis with no new analyses shown. Arguably it could / should precede Section 6 for reader comprehension. It could also be reduced substantially, if the Editor is concerned regarding overall paper length, without impacting the ability of the reader to understand the paper as a whole.

5. Section 8 should be recast as an overall synthesis of results section that helps the reader bring together the various strands of evidence and outlines the case being made. This may be the appropriate location to discuss impacts on temperature, precipitation and other societally relevant variables if the authors prefer to arise that here (see comments on Flato et al in the second section discussing the received short comments and author responses), associated with current Section 8.2.

   Such a revamped synthesis section is where, logically, a revamped and expanded Section 10.2 (see comments elsewhere) would also sit at the end of, leaving current Section 10.1 – suitably modified - to be a conclusion section in its own right.

6. Where cross-referencing is done it should always cross-reference the specific section or sub-section being considered. The only exception to an enumerated (sub-)section pointer should be ‘in the previous/next (sub-)section’. Otherwise forward / backward references should always be to the specific (sub-)section in question to aid the reader rather than a vague ‘as shall be discussed / returned to later’, which is entirely meaningless to the reader. Furthermore, such cross-referencing should be limited to essential cases only. Please consider carefully whether forward / backward references are essential to make life as easy as possible for the reader.

Figures

The figures all use a single color schema regardless of the geophysical variable being considered.

Firstly, this color scheme is not color-blind friendly, which serves to reduce accessibility for the not inconsiderable proportion of the population who are color-blind. Several sites exist such as colorbrewer2.org which highlight color-blind safe schema.

Secondly, thought should be given to appropriate color scales and color ordering. For example, precipitation should be brown to blue or green – not blue to red. The current blue=dry to red=wet schema runs entirely counter to societal expectations and puts an entirely unnecessary interpretative burden upon the reader.

With some thought regarding color schema the graphics can be made substantially more accessible both scientifically and to color-blind readers. See Chapter 2 figures in Plate 2.1 in any of past several years State of the Climate reports in BAMS for an example of better ways to do this than is the case in the current manuscript, that provides appropriate color-schema for all variables discussed in the manuscript.

Finally, please also provide the units (e.g. mm/day, K, Wm-2 etc.) under each and every colorbar to aid reader interpretation.

Original reviews and responses thereto

In this re-review I have concentrated upon solely that subset of 'Short Comment' reviews arising from recognised scientific experts. I have tried to assess the review, the response, and any modifications that were made in the revision. I am assuming that the other reviewer shall assess whether the authors were sufficiently responsive to their own review. The lack of traditional responses, alluded to previously, has made this a far harder task than it should have been.

Herein I am largely picking out those issues, which I see as still open, and of sufficient import that they require further addressing in subsequent revisions. To be clear, for the editor, many of the points raised were addressed in replies and / or revisions, and I do not highlight these here for expediency. I also do not re-raise points already substantively dealt with above.

Drijfhout et al comment

Beyond the noted issue over the gross inappropriateness of aspects of the response, and noting that Drijfhout et al share my concern with regards to editorializing, I believe that the following aspects have not been sufficiently addressed:

1. That the Eemian cannot be directly compared to any future climate eventuality.

   Noting that the authors reject that assertion in their response to the Short Comment, I nevertheless have sympathy with the reviewers' point here. The unspoken implication through the use of the Eemian paleo-evidence is that it offers some sort of meaningful analogy to what may occur to support the modelling exercise. Otherwise, why include it, and why does it build evidence per the letter to the editors? The authors cannot have their cake and eat it here. Either it is there because it is a useful analogy, or it is an unnecessary distraction. Which is it? It can't be both. If it is there as an analogy, then it needs to be made more explicit in the final manuscript how adequate an analogy it may be. This may be best achieved by elevating elements of Sections 6.4 and 7 to early in current Section 6.

   In my view, upon balance, it is required to more carefully and explicitly caveat against such a direct implication being drawn by the unwary reader in the opening section of the paleo discussion. This would ensure readers are explicitly aware that they cannot and should not imply that the 5e stage can be used as a direct analogy to what may happen with increasing GHG burdens in the 21st Century. The distinctions between 5e viz. (at least): i) GHG burdens, ii) seasonality of solar radiation receipt, and iii) potentially ice-sheet configuration, should be stated categorically up-front at the start of the present Section 6 to ensure that the reader has all information necessary to interpret what follows appropriately.

Specifically, I would note in regard to this point, that re-reading the paper highlighted more strongly to me how distinct the solar forcing seasonal cycle at many latitudes was in both hemispheres during 5e in comparison to present day. Given the potential import of seasonality of radiation receipt to the posited ocean heat content / ice sheet dynamics / responses and their hemispheric asymmetry, it is not clear to me how appropriate an analogue 5e could be to what may happen in the 21st Century which will be dominated by a more globally homogeneous LW forcing perturbation. Differences in seasonality of forcing during 5e are an order of magnitude larger than anything we would see under GHGs, even under the most pessimistic assumptions about our collective political and societal responses to the challenges we are undoubtedly confronted with.

Therefore, it is unclear how much confidence the scant paleo-evidence provides to the modelling results given the distinct forcing mechanisms at play. I believe that discussion of this aspect is required within the manuscript when introducing the Eemian section, to make explicit to the reader the potential limitations of its use as an explicit analogue.

2. The sea-level multi-stage issue they highlight is in my view important for understanding (or trying to) underlying processes. It likely does matter how many relative maxima there were over the Eemian, as this likely says something about mechanisms and/or stability of ice sheet collapse and regrowth, which may logically imply something about how similar the ice sheets were to today's configuration. The reviewers did not dispute the presence of the final maximum, but I think their point here warrants further attention and explicit discussion in the paper.

3. The reviewers' point about the plausibility of the high rates of freshwater discharge imposed in the model experiments is well made, but was not responded to. You can force your model with any forcing you want, obviously. However, the experiment is only going to be useful and informative if the prescribed forcing is possible to attain in the real-world. This point was also variously raised by Flato et al., Pelto and others in addition. It is clearly a fairly broadly held criticism within the expert community, that requires more substantive discussion and caveating in the manuscript than is currently the case.

   I share the reviewers' concerns that the high rates assumed may be entirely implausible. If the authors wish to make inferences based upon such high discharge rates, it is necessary to at least discuss more fully mechanistically how such a high rate could be attained and then maintained. From where, specifically, is the ice coming and what physically is enabling such a freshwater discharge growth profile in the real-world? I see some discussion of this but that discussion is one-sided and does not reflect the counter-views given by relevant experts in several Short Comments.

If the prescribed freshwater forcing is unrealistic then the results have limited real-world utility, so it is necessary to address more fully this legitimate concern, and perhaps remove runs for which it is felt that the rate of discharge implied is implausible given known process understanding.

4. The reviewers' comment about the intensity and frequency of storm tracks was well made but not responded to. Further discussion of this aspect in a revision and a response to the point raised by the reviewers is warranted here. Note that Michael Wehner made additional comments in this regard which were not addressed in the response to his comment which pointed to a non-responsive reply elsewhere.

Comment arising from Dale Berner

1. I would like to see many of the recommendations for future work mentioned in this review highlighted in Section 10 where you discuss future possible work directions. Most of these seem ostensibly sensible but I would welcome replies to them that make clear which you concur are sensible or not and why. In general what is now Section 10.2 is still too light on detail. I'd like to see much more specific recommendations spelt out.

Comment arising from Matt Whipple

1. I am unconvinced by the response to the points raised regarding the Greenland ice sheet by the reviewer. I think it is likely to be important to understand more fully what proportion of the contribution to the 5e sea-level maximum was from Greenland and Antarctica respectively. I don't feel that the response or the revised paper adequately deals with the issue.

   The point about relative Greenland stability in the Eemian raised in the response: i) is contradicted by the comment from Jason Box regarding possible Greenland mechanisms; and ii) contradicts much of the modelling work in the paper that injects substantial amounts of freshwater from Greenland, so implies a substantial degree of instability. Again, the authors cannot have their cake and eat it. Either Greenland is inherently somewhat stable, or it is not. It is not Schrodinger's ice sheet (melting one minute you look at it, stable the next), so the authors cannot argue contradictorily at different points in their manuscript and review comment responses. Greater consistency is needed here, and this likely requires acknowledging the potential Greenland contribution to 5e per the reviewer comment, and a greater discussion of Greenland dynamics generally, with avenues for future work spelt out in a revised current Section 10.2 text.

2. Similarly, the reviewer's comment about the lack of evidence that the WAIS collapsed during the Eemian from available cores calls into question the authors' contention regarding Antarctic discharges, and is not adequately dealt with. At the very least this ambiguity needs to be acknowledged in the paleo-evidence section, and possible avenues of future work to assess whether the WAIS collapsed or not during the Eemian should be added explicitly to the future required work in current section 10.2 to confirm or refute the authors' hypothesis.

Comment arising from Michel de Rougemont

1. While in my view the reviewer overstates their case in their major comment, the overall contention that SLR is not solely a function of Carbon Dioxide is trivially true. Further, the reviewer makes the valid points that: i) we are arguably yet to see a statistically robust acceleration in the available direct observational record; and ii) the SLR is to date dominated by non-ice sheet contributory terms.

   To get the posited SLR raises foreseen in the manuscript would require an extremely rapid acceleration in discharge and melting of ice sheets. As noted elsewhere this is deemed unlikely by many in the community who submitted comments, myself included, on a range of energetic, mechanistic and theoretically based grounds.

   When discussing the observed SLR curve the authors should be much clearer in the revision that most of this rise to date results from non-ice sheet processes – thermal expansion, glaciers, terrestrial storage. Presently the unwary reader may infer that the observed trends arise due mainly to ice-sheet loss, which is unambiguously not the case.

   The authors' contention that observed SLR is accelerating is possible by segmenting quasi-arbitrarily and showing three rates that differ, but this is not equivalent to robustly concluding that the change is a statistically significant changepoint in the series behaviour. That would require a test that is as yet un-run. As the authors' astutely acknowledge in their response – it is important not to fool oneself. If they wish to contend one or more changepoints in SLR behaviour, then it is incumbent on them to prove it using appropriate timeseries changepoint detection techniques readily available in the statistical literature. Arbitrarily segmenting the series without a clear basis and then implying a physical change is not good science.

Comment by Mauri Pelto

1. Mauri Pelto raised a number of suggested specific areas regarding understanding ice-sheet dynamics that should be investigated, and these should be pulled through to Section 10.2. Furthermore, the valid concerns raised about the physical plausibility of sustaining a long-

term doubling regime raised by Mauri Pelto should be raised appropriately when discussing the model experimental set-up and then again in section 10.2.

Comment by Dr. Colgan

1. Dr Colgan correctly identifies discussion of non-linear sea-level rise in AR5, and highlights the relevant sections. Although the authors have acknowledged this in the redraft, they have done so in a way that in my view is not necessarily objective and still leaves an unduly negative impression as to what was actually stated within AR5. I would like to see a more explicit acknowledgement of what was discussed and why it was precluded in the summary figures, out of fairness to the AR5 authors. It is valid to state this is a weakness in AR5 numbers, but to be fair also requires pointing out why these processes were omitted in the final numbers despite being assessed.

Comment by Greg Flato and colleagues

1. I was particularly struck, in the Short Comment by Flato and colleagues, by their point upon the very large changes in temperature that accompany the hosing experiments. It is, indeed, key that this be highlighted more strongly, and I believe further efforts are warranted in this direction. But, this yields also an obvious further question as to what may happen to other societally relevant parameters such as rainfall.

   It would, perhaps, be wise to add a short section to the modelling results discussion or the synthesis section 8, to explicitly assess the changes that would arise in the event that the model runs were realised in the real-world. Surely, in terms of societal relevance, it may well be the global temperature and precipitation response that would be the largest impact upon society as a whole, at least in the medium term (decades hence), rather than the sea-level response per se? The paper concentrates upon SLR, and largely ignores that there are very substantial changes to our current expectations for additional societally relevant aspects of the climate system in the experimental results. These may have more broad-reaching impacts. These results should be shown and discussed.

2. The use of a -15C input of water is clearly unphysical (it would immediately freeze again) and likely to yield issues with model vertical mixing; a mixing that is already far from perfect. I did not feel that the response or revisions at this point adequately addressed this point. The lack of permissible feedbacks in the model is also a limitation. Further caveats are required to this end when describing the model experimental set-up and results. Basically, as well as model limitations which are already reasonably articulated, there are

arguably experimental design limitations which relate to realism both of the freshwater forcing applied, and the mechanism by which it is realised. These caveats need to be bought out more strongly in the manuscript.

Response to reviewers and letter to editors

It is regrettable that the document on the public record is the response to reviewers, which constitutes primarily a commentary on the process rather than a substantive response.

The end of the response to reviewers is particularly unfortunate in its editorialising on the current reviewer's position.

a. First, as noted above I am not trying to obfuscate anything, and that accusation is both baseless and offensive. The authors themselves state they do not expect policy makers to read papers and figure it out, so they destroy their own case here. The scientific literature in general, and ACP for certain given its remit, are not the appropriate places for discursive sweeping policy statements. I strongly maintain that they should be removed. Peer review provides an imprimatur. It effectively states that the journal and the peers who reviewed the piece broadly agree with the statements given. As such, I cannot accept the counter-contention given that it is valid for the authors to insist on inclusion of policy statements.
b. Secondly, and more importantly, I am not representing any positions other than my own, based upon my own scientific knowledge and experiences, which are far, far, broader than drafting aspects of IPCC AR5. Trying to paint the review as being in some nominal sense an 'IPCC' review or representing an IPCC position (whatever that is) is unwarranted and unhelpful. I undertook the review (and this re-review) in an entirely personal capacity - not for, or on behalf of, IPCC. That I have contributed to IPCC in the past is entirely tangential and has no bearing here.

Accusing the scientific community in the response on record of group-think is something I would expect of certain, more hysterical, quarters of the blogosphere. It does the authors no favours being associated with such a statement. What is the basis for this assertion? Nowhere is the supposed group-think being challenged actually spelt out in the response to reviewers, so the whole passage is utterly meaningless as a result. It seems baseless to me.

If it is about sea-level findings in AR5, then this is both at odds with the recognition of the discussion of non-linear effects in AR5 acknowledged elsewhere, and ignorant to the wealth of information provided by e.g. Dale Berner outlining substantive literature which gives already potentially higher values than shown in AR5, some of which was assessed in AR5. That breadth of

literature and discussion of SLR does not strike me as being a community in the thrall of some mis-placed group-think on the issue. It strikes me as a difficult problem that many groups are trying to solve. As evidenced by even more recent papers that have gained substantial media attention in recent weeks on the subject of e.g. Antarctic ice-sheet vulnerabilities and associated SLR contributions, there is still substantial work on this aspect on-going.

The paper does depend, as noted in my review and others, on several lines of evidence. I will concur and concede that these may in some cases be less a chain and more an inter-linked series of strands and apologise for the mis-representation of such. However, regardless of the exact tautology, the fact stands that the result is dependent upon all the distinct aspects pulling together to tell a coherent story. The case, therefore, remains that if one strand is the wrong, the case for the whole is substantively diminished. This then directly relates to the major comment about the unwarranted degree of certainty being communicated in the title, abstract and conclusions. All strands considered have copious caveats that preclude making definitive conclusions.

Contrary to the response to reviewers on record, the letter to the editors is more measured, and contains more specifics. It would still have been incredibly useful, however, to have a point-by-point response to the two formal reviews. Some observations on this document (limited to those very few aspects not already covered previously) that may help the editor in coming to a determination are given below:

1. I would concur that the paper does not need, and would not benefit from, being split asunder and considered as several interlinked contributions.

2. I would maintain that showing the same modelling experiments for a different, substantively independent, model would help us understand the confidence warranted in the modelling results. Without doing this substantial caveats to the model analysis are *de facto* required for the piece to be acceptable. In particular the recognised inadequacies in the control sea-ice and AMOC initiation are major issues in interpreting more broadly from the current single model approach.

3. I agree 100% that the 2C being safe is an extraordinary claim that requires extraordinary evidence. As stated in my original review I see the whole 2C framing as highly disingenuous and dangerous. However, I think there was some unfortunate and accidental mis-interpretation of my intent here. Flato et al said it better than I did in their review, but I shall expand still further here.

   If the findings in the manuscript are true, then it implies every single shred of adaptation planning being undertaken based upon current CMIP5 model runs and associated RCMs and statistical downscaling is mis-placed, and that even the sign of the temperature change we are planning for is wrong. This would require truly extraordinary changes to major planning, and is a truly extraordinary paradigm shift in what we

should be doing as a global community to respond to the threat of human-induced climate change. This has all the hallmarks of an extraordinary claim to me and, as such, I would maintain requires extraordinary evidence.

4. As alluded to elsewhere, I do find the Section 10.2 useful, but believe it could benefit from very substantial further expansion and specificity, including input both from several of the Short Comments and from insights arising from the entire author team.

I'd expect an adequately redrafted Section 10.2 to cover at a minimum 2-3 pages and cover many of the specificities raised in many of the Short Comments. It should be limited solely and exclusively to avenues of scientific investigation necessary to confirm or refute the authors stated contention. Discussion of geoengineering options etc. would be seen as non-responsive here.

Review of the resubmitted manuscript

This section raises solely specific points that either are not addressed above, or where specific suggestions are warranted. The points are raised in the order in which they arise within the manuscript, rather than in any order of importance. Their importance should be easily inferred from the comment itself.

1. The abstract should highlight the temperature response in the model experiments and possibly the precipitation response. [Other abstract comments covered in above points]

2. Line 75, why the need to highlight this wasn't cited in prior IPCC reports? Very many things weren't. It doesn't mean it wasn't considered. Many papers are considered but not eventually cited. This seems unduly antagonistic towards IPCC and I see no rationale for its retention. It does not really help the reader understand or interpret the paper. Please delete.

3. Line 81, why remarkable paleodata? Remove remarkable which is a value-laden assertion and at odds with more circumspect language elsewhere.

4. Line 82, add appropriate caveat here to be clear about implications about potential future sea level and storms.

5. Lines 85-88 need to clarify here that these assumed rates may or may not be realisable in reality, per several Short Comments received.

6. Line 91 – what IPCC-like? The modelling group that designs these experiments is the Coupled Modelling Intercomparison Project – CMIP (http://cmip-pcmdi.llnl.gov/) – so revise as CMIP-like so that your text appropriately gives credit to the actual body that oversees these experiments. They, like IPCC, rely upon broadly volunteer effort to design and implement the modelling strategies, and they should be recognised for such.

7. Line 105 – as above these were CMIP modelling studies so could be referenced as CMIP3 / CMIP5 and the CMIP3 and CMIP5 overview papers cited instead of IPCC reports.

8. Line 119 – correctly seems too strong here – better orients – or similar would reflect the presumed remaining uncertainty?

9. Line 126-127 – the D= term makes no logical sense and I assume the parentheses are mis-prescribed. Otherwise why include the 1000m at all, because that condition can never be met as this term is currently formulated?

10. Figures 1-3 – please provide units under the colorbars rather than, or instead of, in the title. Use of more appropriate and color-blind friendly schema (see e.g. colorbrewer2.org) that are distinct for different parameters would serve to aid reader interpretation. See major comment in first section.

11. Figure S2 - precipitation should be a more appropriate color-schema or at an absolute minimum the color-scheme should be inverted so bluer hues imply more and not less rainfall. Ideally a brown to green or blue schema that is color-blind friendly should be used. SLP should use a further distinct color-set so it is clear to the reader that the three columns are for distinct parameters.

12. Line 168-170 requires a supporting reference or deletion.

13. Figure S3 – please use a more appropriate color-schema in the sea-ice column because the current schema implies ice loss to the unwary reader for the red end hues when in reality it is very substantial ice cover. NSIDC uses an appropriate grey-blue scale schema or the Uni-Bremen group similarly uses a commonly seen scheme that readers may expect to see. For cloud please use a blue (low %age) to grey (high %age) schema.

14. Figure S4 is of limited utility without in addition indicating on the figure by e.g. horizontal lines what the modern observed fractions are, to be able to ascertain how large the biases are.

15. Line 192-193 rather than being qualitative here, presumably whether the effect is statistically significant could be calculated and reported instead?

16. Lines 207-219 need recasting to reflect that the doubling times may be hard to sustain etc., per the copious comments discussed in the earlier portion of the review.

17. Lines 220-225 should reflect the numerous comments given in the main review sections above.

18. Lines 228-229 are not needed. Instead such text should be bought out to an opening paragraph at the start of Section 3, that outlines the structure of the section as a whole, or be pulled into the suggested paper outline section of a revamped introduction.

19. Line 237 – is it injected on or injected into?

20. Lines 237-241 should caveat that feedbacks between forcing and response are omitted, per the comment of Flato et al.

21. Figure 8 please use a blue to grey scale for the cloud cover panels to enable ease of reader interpretation. Please place units key under each colorbar.

22. Figure 9 please place units under colorbars

23. Lines 284-286 should be deleted or moved up into early section introduction that outlines to the reader the overall structure of the section.

24. Lines 289-294 should be in the revamped introduction rather than in this section.

25. Line 314 – what IPCC studies? IPCC does not undertake studies and indeed is not directly permitted to. Its purpose is to undertake an assessment of existing literature and state of knowledge. Thus delete 'based on IPCC studies' here, which is both meaningless and untrue.

26. Figure S9. Please use a more appropriate color scheme and place units under the colorbars. Greater rainfall should be in a blue or green hue and reduced in a brown hue per societal expectations.

27. Line 375-377 – to what extent is this response lag questionable given the earlier analysis of control showing 500 years spin-up required to create deep AMOC circulation? This may imply that results regarding resumption of AMOC throughout Section 3 and Section 4 experiments are questionable and this caveat needs to be made more apparent here and elsewhere. This reflects the need to verify using an entirely independent model alluded to both in my original review and above to build confidence, which needs expanding in Section 10.2 compared to present.

28. Figure S10 - please place units under the colorbars. Consider elevating to main text per earlier segments of the review, because the geographical distribution of the temperature change is an important aspect, rather than just the global-mean behaviour as given in Figure 11.

29. Figure S11 – error in caption. I presume should refer to Figure S10?

30. Figure S12 – error in caption. I assume should refer to figure S10 or Figure S11?

31. Line 399-401 should note that such injection lacks physical realism, as at such cold temperatures the water would instantaneously freeze. The model can treat it as liquid but in the real-world, to my knowledge, even for highly saline water (which it isn't) it is frozen. Just because you can do this in the model doesn't mean it can happen in the real-world.

32. Line 428-430 – again, this NADW response may arise from model issues alluded to in the control run discussion, rather than reflect a real-world response lag.

33. Line 435-441 – these forward references don't really aid comprehension here, and would aid readability if removed. Please delete or modify.

34. Lines 444-447 – as previous comment. Inclusion of this text adds very little value here to interpreting the currently discussed results.

35. Line 455 – given that this section is discussing primarily the evidence for advanced freshwater injection the title should include something like 'Observational basis' to reflect this.

36. Line 464 – flux into the Ocean?

37. Line 489 – use of qualifier 'remarkable' is value laden and should be removed.

38. Line 539 – what does (S20) refer to?

39. Line 569-570 – it still requires also a source that is susceptible to such doubling – see review comments in particular from Mauri Pelto. A much stronger caveat is warranted here about whether such doubling is attainable in the real-world.

40. Line 586-587 – Section referencing is mixed up here. Please correct. Also, this statement is stronger than the relevant sections suggest is the case – there are residual uncertainties in both.

41. Line 599 – does this figure arise from the authors, or is one or more sources warranted to be cited within the figure caption?

42. Line 610 – constitute rather than stimulate.

43. Line 629 – how do models constitute a 'paleo-affirmation' exactly? This seems very mixed up and requires clarification in redrafting.

44. Lines 642-643 are not necessary and should either go in a whole section introductory paragraph, or be deleted.

45. Lines 644-646 then 646-656 – it seems odd to explain the first factor without indicating the second factor first. If you are going to start by saying there are two factors each should be introduced before you deep dive into an explanation of the first. So, rephrase this or bring up a brief intro of the second factor at the end of the second sentence here.

46. Figure S15 - please put units under each colorbar, and consider using a range of available relevant color-blind friendly schema to help readers differentiate that distinct panels refer in some cases to distinct physical facets of the climate system.

47. Line 655 – why underline under? It is probably not commensurate with journal style guidelines. But, regardless, I don't see it as appropriate.

48. If you are going to claim that the trends apparent in Figure S17 are similar to observations then Figure S17 should include values plotted from one or more reanalyses products such as ERA-Interim or JRA-55 to support that contention, rather than requiring the reader accept it as an article of faith.

49. Line 706 – surely there is a more scientifically appropriate terminology here that describes the process better than 'wrenches'?

50. Line 708-709 Sentence is not needed. Delete.

51. Line 737-743 – please modify to acknowledge that IPCC explicitly did not include the semi-empirical estimates in their final assessment because they assessed to have low confidence in them. This would be a fairer reflection of the underlying assessment process.

52. Line 755-757 – this is an assertion that requires a robust underpinning statistical analysis to be verified and remain.

53. Line 757 and Figure S19. Given that the two analyses are not measuring exactly the same thing we'd expect them to differ in the details. There is also no discussion of the figure. Both the sentence and the supplementary figure should be deleted, as they add no value.

54. Line 758-760 - omits to mention that the vast majority of the global SLR observed to date arises from non-ice sheet processes. This should be stated explicitly and referenced appropriately.

55. Line 769-770 - the observations are also consistent with a range of alternative doubling times given the self-evident annual to multi-annual variability, and so it should be stated explicitly here that the observations are also consistent with a number of much longer doubling periods, as well as explanations much more complex than simply a doubling rate.

56. Line 784-785 it is questionable whether the modelling results should be included in the observations section. Delete from 'but this …' on to avoid conflation between a purely observational analysis and model results.

57. Line 825-826 - While the record is marginally consistent with a decadal rate curve, it is also consistent with much longer doubling times or more complex behaviours, and this should be stated. By the same token, it appears inconsistent with anything faster. Furthermore, the review of Mauri Pelto raised concerns over the realism of such doubling assumptions, which should be caveated here.

58. Lines 827-834 - It is unclear whether this paragraph relates to Greenland, Antarctica or both. Please specify by appropriate modifications.

59. Lines 835-838 should be modified to make clear that the worst case scenario is multimetre rise, but that the observations would also be consistent with much lower overall contributions. Only additional data shall clarify. It is not tenable to concentrate solely on the extreme tail, without acknowledging that the data are also consistent with far less catastrophic outcomes.

60. Lines 838-844 should be directly associated with the Greenland discussion in lines 801-826, rather than placed where they currently are, for reader continuity.

61. Lines 850-872 have nothing to do with observations and as such arguably must be moved elsewhere, after taking into account the comments on specific passages given below.

62. Line 850 CMIP3/5 models and not IPCC models

63. Line 852 As previous comment

64. Line 853 – your model is a modification of a CMIP5 model so it would be more correct to talk about your model experiments, which avoids the impression that your model is entirely new – it is not – it contributed runs to CMIP5, which were considered in IPCC AR5. This needs to be corrected to avoid the unwary reader from potentially misinterpreting here.

65. Figure 22 – the low variance prior to 1980 in the right hand panel is an artefact of processing choices in the SST algorithm used. Specifically, ERSSTv4 uses HadISST sea-ice and prior to 1979 this is a repeating climatology and then satellite observed sea-ice cover thereafter. This has the direct effect of greatly, and artificially, reducing variance in SSTs in the Southern Ocean prior to 1979. This should be noted or, preferably, the observed series shown only from 1980 onwards. See Huang et al., 2015 or Huang et al., accepted, J. Clim for details.

66. Figure 22 – please update both panels to run through 2015, which should be possible to calculate and include on the timescale of any resubmission, and ensure it is up to date. Please modify discussion accordingly if required.

67. Line 860 – see the above comments regarding how you are referring to these models.

68. Line 861 – model experiment and not model

69. Figure 23 – can observations be added to the left hand panel? For the right hand panel please update through 2015, when ice cover returned to the long-term mean, and discuss accordingly regarding the possibility that there may be a mismatch between modelled and observed variance in the parameter, rather than necessarily a difference in timing of emergence of a trend. Currently either interpretation is plausible, and you need to acknowledge this. Only more years of record could cleanly differentiate the two possibilities.

70. Lines 884-889 – by the same token, the growth of sea-ice may have resulted from the string of La Nina type conditions that in part led to the much argued warming 'hiatus' globally, that has almost certainly now stopped. That the sea-ice growth in recent years may have resulted from a string of La Nina events behaviour should be noted here if the authors wish at the same time to invoke El Nino to explain the return to normal conditions in 2015. The authors cannot posit an ENSO response without acknowledging the potential logical interpretation that follows as to why ice grew in recent years in the first place. Again, this is arguably a case of the authors wanting to both have their cake and eat it.

71. Figure 24 left hand panel should include the 8 years of good observations from the transect. Note that the NOC team recently returned, and additional years of data are likely available upon request as a result.

72. Figure 24 right hand panel - please update with SSTs through 2015 to reflect the latest data which may impact interpretation.

73. Lines 900-903 – the purported slowdown returning in recent years is not supported in my interpretation of the Figure 24 right hand panel, which shows enhanced interannual variability but no apparent trend. The panel also highlights that the model under-estimates the inter-annual variability, and this should be highlighted. Particularly so as ERSSTv4 itself is overly smooth owing to the EOT smoothing (Huang et al., accepted, J. Clim).

74. Line 905 – talking about an event being achieved seems odd. Please rephrase in a more scientifically appropriate manner.

75. Lines 908-929 belong in Section 6, as they are discussing paleo evidence. Please move to an appropriate position within that section and delete here.

76. Line 934 – insert qualifier 'likely' before exerts, to reflect the uncertainties more properly here.

77. Lines 951 to 953 should be moved to what is now Section 10.2 and expanded accordingly.

78. Lines 955-959 are not relevant to the sub-section being discussed, and should be deleted or moved.

79. Line 971 – citation of Masson-Delmotte et al. 2013, which is the AR5 chapter, raises a point regarding the varied way in which IPCC is being cited. Please check all citations to IPCC. Citations to chapters and summary materials should follow the referencing guidance given at http://www.climatechange2013.org/images/uploads/WGIAR5_Citations_FinalRev1.pdf. The reference here is correct, and it is the remaining references to AR5 that need to be modified accordingly. Please search for and change all other IPCC references so they are ({appropriate-surname} et al., 2013), and update references accordingly.

80. Line 972-974 – not necessarily. If the sea-level response in the Eemian was a response to seasonal changes in solar forcing we would not expect the same sea-level response to potential GHG forcing in the coming Century. This relates to major point arising from the KNMI-led collaborative comment discussed in the second segment of this review.

81. Line 1031 – should there be a + in front of 3-4m?

82. Line 1033-1036. First sentence is repetition from earlier in section. Second sentence adds no interpretative value to the section. Therefore please delete this segment starting End-Eemian forwards.

83. Line 1164-1177 – Why even mention the boulders here, as this sub-section is not about them? If this paragraph is summarising the preceding sub-sections then please make it its own sub-section entitled something like 'Summary of evidence from Bahamas and Bermuda' or similar. Otherwise this whole paragraph feels out of place.

84. Line 1178-1179 – this should either be in the main introduction, a section introduction that tells the reader what to expect in the section as a whole, or deleted.

85. Line 1184 – global glacial conditions. Geologically speaking, we remain in an ice age today because we still have two substantial land ice sheets, as the authors are repeatedly alluding to throughout the text!

86. Line 1189 – delete 'as discussed in the next section' as it adds no interpretative value to the reader here.

87. Line 1197-1200 it may be worth being clear why CH4 is used – presumably because it is globally well mixed but has a relatively short lifetime that allows annual-scale or at least decadal-scale gradients to be resolved?

88. Line 1202-1205 please state what caveats a few decades uncertainty in their synchronisation may arise for your subsequent analysis in this section at this juncture in the interests of full disclosure.

89. Lines 1242 to 1253 arguably should be in the later Section 7.4 discussion of D-O events, or cross-referenced to there. The two sections should be reconciled to ensure against repetition and improve the overall paper flow.

90. Lines 1322 to 1329 are a key facet that should come much earlier within Section 6. Indeed, arguably Section 6 would make more sense if Section 6.4 were made Section 6.1. It makes little sense to give this scene-setting section as an afterthought to the section as a whole. The reader would be aided by its coming first and it would help to naturally address major concerns raised about providing caveats about ability to use the Eemian as a direct analogue. I strongly recommend that Section 6.4 should come first within Section 6 and set the scene for the remainder of the section.

91. Lines 1330-1333 are not required. Either move to start of Section 7 or delete. The reader does not repeatedly need short segments that tell them what comes after the next section heading. The paper is already long – don't make it unnecessarily longer.

92. As noted in the major comments in the first section of my review, it is unclear to me what the distinguishing feature of Section 7 is from section 6 and, therefore, why a section break is warranted.

93. Lines 1345-1353 are broadly repetition of Section 6.4 text. Please reconcile and discuss just once.

94. Section 7.1 feels mainly like a section introduction but fails to then outline what shall follow, so is demonstrably incomplete if that is the intent here. If that is not the intent its not entirely clear what the purpose of this section is to the paper as a whole.

95. Line 1459 increasing? increases? Regardless, increase is not grammatically correct here.

96. Lines 1463-1468 should make clearer the complexity in the Carbon cycle, and that some of the Carbon is removed quickly, rather than stressing solely the fact that some remains for c.100kyr.

97. Line 1469 – please reference where you suggest this. If it is here then the suggestion is not particularly well justified. If the suggestion arises elsewhere the relevant sub-section should be referenced.

98. Lines 1515-1517 – this finding should be picked up again in the expanded Section 10.2 and potential research to address it discussed.

99. Lines 1520-1545 – this text needs to acknowledge inevitable uncertainty arising from dating issues that may impact the interpretation of the event, through an appropriate caveat or caveats.

100. Line 1587 – a definitive statement that it could not seems unduly certain. There could have been changes in albedo that allowed surface melt per earlier discussion by the authors and Short Comment by Jason Box. We simply don't know. Deemed unlikely or similar language would better reflect inherent limitations of how certain we can be here, given the paucity of direct evidence available to work with.

101. Line 1590-1591 – likewise, this characterisation is too definitive and should be couched in more appropriate language, that recognises there may exist alternative explanations.

102. Lines 1595-1598 are repetition of earlier text. As noted earlier I suggest these points be moved up to start of current Section 6 to ensure proper reader interpretation of the evidence.

103. Lines 1601-1604 should be moved to current Section 10.2 and expanded.

104. Line 1628 – CMIP and not IPCC.

105. Lines 1633-1640 – these should be expanded upon in current Section 10.2. Lines 1639-1640 should probably be moved there.

106. Line 1641 – please cross-reference back to the section(s) and / or figure(s) where this was shown in addition to referencing the SI figure. If it is only shown in SI and it is a key assumption then it should in all likelihood be elevated to the main text and discussed further here.

107. Lines 1648-1649 ignores that the modern injection, were it to occur, may be in both hemispheres whereas the 5e injection by all accounts given by the authors was a SH only injection (although note contention on this point from some short comments discussed above). This needs to be stated here to enable proper interpretation by the reader. The responses may well be different per Section 3 and Section 4 analyses.

108. Line 1651 – This section title is disingenuous. The section details model run results (from Section 3 or Section 4 is unclear), using the hosing experiments. It is therefore not correct to title this Eemian storms, which gives the impression it is direct analysis of Eemian storminess – it is not. Suggest – Modelling insights on Eemian storminess using freshwater injection experiments - or similar which more adequately reflects what the section is actually about.

109. Line 1652. Please start this section by being explicit as to which of the myriad modelling experiments you are discussing results from here, and whether freshwater injection is in both hemispheres and on what doubling rate.

110. Line 1652 – increases simulated sea level pressure …

111. Line 1653 – increased rather than added

112. Line 1654 – strong -> stronger and please quantify the impact and report it here (move up text from lines 1669-1674).

113. Line 1658 appropriate -> necessary

114. Figure 32 – see major comment about appropriate colorscales and key labels

115. Figure S22 – See prior comment. Please also clarify in the figure caption what the numbers top right of each panel refer to.

116. Line 1669-1670 – make clear that this is in your model simulations.

117. Line 1675 – how is this assessed to be robust? If you have not assessed whether it is significant, please remove the value laden robust here which implies such an analysis has been performed. If you have undertaken such an analysis that underpins this statement please outline the result quantitatively.

118. Line 1683-1684 – is this effect included in the model or not? If it is not then it should be stated so and possible future work to address this shortcoming should be discussed in the expanded Section 10.2 text.

119. In Section 8.2 please clarify by appropriate referencing to Sections 3 and / or 4 which particular modelling runs you are considering. It would arguably be valuable to discuss results for all three doubling rate experiments plus the more 'standard' runs here, and that would add undoubted value to the reader's interpretation here.

120. There appears to be a disconnect between section 8.1 discussing sub-tropical impacts, and section 8.2 that discusses solely mid-latitude impacts. At a minimum, section 8.2 should remark upon whether the Eemian type response for the sub-tropical locations also exists in the projection runs with freshwater hosing being discussed in Section 8.2 for narrative continuity.

121. Line 1713 causes an increase (missing an)

122. Lines 1772-1774 are a truism and arguably not necessary.

123. Line 1774 (if retained) communities and not community.

124. Lines 1783-1784 are overstating the authors' case in my opinion and require caveating.

125. Line 1787-1790 – I'm not sure that the correct way to test a hypothesis is to try to confirm it unless we wish to enter the realm of post-normal science. The scientific norm is to test a hypothesis by trying to disprove it!

126. Line 1791 – CMIP not IPCC. Also note that your model is one of the mis-named IPCC models here in that GISS-ER contributed to CMIP5. This, therefore, needs to be rewritten accordingly. It must be made clear that you are running a sensitivity set of experiments with a model that submitted in pretty close to the same set-up to the CMIP5 experiments. It is not a new or independent model, rather it constitutes a novel set of experiments that assesses sensitivity to several assumptions / possible permutations in ice-sheet responses. It needs to be couched as such here.

127. Line 1792-1793 With that assumption, we predict the following potential consequences, which warrant further investigation and confirmation or refutation: … - this expanded opening would more fairly reflect the uncertainties recognised elsewhere in the paper. I would find this statement hard to accept without such a modification, as it is too certain otherwise (see major comment in opening remarks).

128. Lines 1812 – 1828 are policy discussion, and not a discussion of what further study is required to confirm or refute your findings. Please delete these from this section and the paper.