## General comments

Khabbazan and Held's paper highlights the nuances which must be considered when using a one box energy balance model for climate projections (the form they focus on is the one presented by Petschner-Held (1999), herein PH99, but any similar one-box model would exhibit the same behaviour). Their major conclusion is captured in the last paragraph of the paper, specifically that calibrating PH99 is 'much more involved than previously assumed' and hence 'future users should carefully consider whether they actually want to use PH99, or whether they prefer a less parsimonious solution'. On top of this, they also present a lovely bit of analysis which shows why a one box model must use a lower ECS than a two-box model if the two are going to respond similarly to a strong mitigation radiative forcing scenario over an ~200 year timescale.

For final proof reading I have included a number of technical corrections below.

Nonetheless, the paper presents some very interesting and pertinent results and I would recommend it for publication.

## Major concerns

## Specific comments

## Technical corrections

page 1, line 12: 'compatible with a maximum global warming of 2 K' → 'resulting in a maximum global warming of approximately 2 K' (as some models project warming greater than 2K for RCP2.6, saying that it's 'compatible with a maximum global warming of 2 K' is too strong, something you implicitly recognise on page 4, line 3)

page 1, line 16: "their" → "the AOGCM's"

page 1, line 17-20: Delete the sentence 'Accordingly, we offer a method to re-interpret already published works based on the 1-box model.' and make the next sentence 'Results that are based on the 1-box model and have already been published are still just as informative as intended by their respective authors; however, they should be re-interpreted as being influenced by a larger climate response to forcing than intended.'

page 1, line 27: 'project' → 'examine' (?)

page 2, line 8: 'deviations' → 'with deviations'

page 2, line 32: 'temperature equation' → 'temperature response to radiative forcing' (?)

page 2, line 33: 'Hereby' → 'Hereafter' or 'In this article,'

page 2, line 34: ‘Quite the contrary’ → ‘Furthermore’

page 3, line 3: ‘Hence in the policy domain, a difference in terms of 0.5 K does matter.’ → ‘In the policy domain, a difference of 0.5 K matters.’

page 3, line 3: ‘In fact we believe that further validation is both necessary and possible at a higher level of consistency’ → ‘[New paragraph] We believe that further validation of PH99 is necessary and possible, at a higher level of consistency than has been performed previously.’

page 3, line 6: ‘2\degree’ → ‘2 K’ (to be consistent with the rest of the article)

page 3, line 7: ‘displaying’ → ‘as these scenarios display’

page 3, line 10: ‘2\degree’ → ‘2 K’ (to be consistent with the rest of the article or change all the K to \degree everywhere)

page 3, line 11: ‘compared to the ECS’ → ‘than the diagnosed ECS’

page 3, line 16: ‘might require’ → ‘requires’

page 3, line 20: ‘Hence’ → ‘To resolve this,’

page 4, line 2: For reasons that aren’t that clear to me you’ve written ‘Sect.’ to mean section in some places and ‘Section’ in others (e.g. page 4, end of line 10). I would guess this will be picked up in copy editing but I think using ‘Section’ (capitalised as it’s a proper noun) is the most common choice.

page 4, line 3: ‘2\degree’ → ‘2 K’ (to be consistent with the rest of the article)

page 4, line 4: delete ‘generically’

page 4, line 5: ‘only RCP2.6’ → ‘only the RCP2.6’

page 4, line 6: ’ and use’ → ‘, use’

page 4, line 7: delete ‘, for the sake of brevity,’

page 4, line 20: ‘section’ → ‘Section’ (proper noun)

page 4, line 20: ‘to then describe’ → ‘and then describes’

page 4, line 21: ‘for a’ → ‘for’

page 5, line 8: degree sign

page 5, line 10: degree sign

page 5, line 22: CO2 should have subscript 2 (and throughout rest of manuscript)

page 7, line 16-19: delete from ‘A proclaimed goal’ until ‘does matter’ (you’ve said it before)

page 10, line 27: ‘After having reviewed their results for our order-of-magnitude estimates of PH99’s accuracy’ I’m not sure I understand this, can you double check? Should it just be, ‘After having reviewed their results’ ?

page 11, line 16: ‘request’ → ‘choose’

page 11, line 18: ‘whereby’ → ‘where’

page 11, line 21: is ‘h’ the best choice of notation for the auxillary function given you use it to mean ‘effective heat capacity’ earlier? Can you choose another letter?

page 12, line 11: degree sign

page 12, line 15: which ‘maximum’? The 2-box’s?

page 13, line 2: ‘can be utilized for any RCP’ → ‘can be utilised for any RCP and the resulting projections are accurate to within X K’ (quantify)

page 13, line 6: delete ‘again presupposing that a 2-box model would emulate an AOGCM qualitatively better than a 1-box model’, you got rid of that statement earlier

page 13, line 10: ‘, which generically ranges from’ → ‘ranging from’

page 13, line 13: ‘any’ → ‘all’

page 13, line 14: ‘would fix’ → ‘fixes’

page 13, line 16: ‘we cannot expect any longer T = T’ → ‘we cannot expect that T=T any longer’

page 13, line 16: ‘to the solution’ → ‘to match the solution’

page 14, line 2: ‘The calibration’ → ‘We find that the calibration’

page 14, line 9: delete ‘in the course of time’

page 14, line 9: ‘Here’ → ‘In van Vuuren et al.’ (or similar)

page 14, line 13: ‘Our article highlights the effects of a naively calibrated PH99 on mitigation scenarios.’ → ‘Our article highlights the effects of naively calibrating PH99 when assessing mitigation scenarios.’

page 14, line 14: ‘However, one should not forget about potential additional mechanisms’ → ‘Additional mechanisms are also possible’

page 14, line 15: ‘mapping on’ → ‘mapping to’

page 14, line 19: delete ‘last but not least’

page 14, line 32: ‘, to a lesser extent, 8.5’ → ‘approximately 0.2K for RCP8.5’

page 15, line 2: degree sign

page 15, line 3: degree sign

page 15, line 7: delete ‘in the rough sense’

page 15, line 30: ‘higher’ → ‘higher by PH99 than by the corresponding AOGCM’

page 16, line 4: degree sign

page 16, line 8: delete ‘assuming that a 2-box model mimics an AOGCM better than a 1-box model’, this assumption is not introduced anymore

page 16, line 15: ‘for which’ → ‘whose’ (reading that again, I think my previous suggestion was a bad one)

page 16, line 20: ‘Hereby’ → ‘Accordingly,’