Response to Referees
gmd-2020-33
Reviewer comments in normal type face. **Response in bold.**

Thank you for taking the time to review this manuscript. We value your feedback and believe that it has led to better, clearer version of the manuscript.

General comments

1. The manuscript describes the emulation of the temperature response to radiative forcing of a simple climate model with an impulse-response function. This is not a novel concept, and corresponds to what is typically one out of a set of many equations in existing simple climate models. The wording used by the authors suggests a more sophisticated, original approach, and is in my opinion a bit misleading. My main objection, however, is that the limitations of the model, which are quite strong, are not sufficiently addressed. The main limitation is that the (climate) model is not really coupled to a carbon cycle component and cannot therefore represent the carbon-cycle-climate feedback. Another important but never mentioned limitation is that the model presented has a fixed climate sensitivity. The approach could be extended by extracting IRFs from models (or model setups) with different sensitivities, but this is not straightforward.

We do believe that this is a new method to incorporate these dynamics into IRF based SCMs (we do not know of other modeling groups that do it this way) but by no means did we intend to suggest that this is the only way it has been done. We've expanded text in the introduction to make it clear to readers that there are several ways to approach this.

We've also added a new section 2.1 Parent Model Description (starting on line 66) to provide a clearer description of the connection between HIRM and Hector; we believe that this addresses your concerns related to the carbon cycle and limitations of the model. Model limitations are also discussed in lines (307-313).

Thank you for pointing out that HIRM could be configured with different climate sensitivities, that is a great idea and we have incorporated it as a future area of research in the discussion section (lines 344 - 346).

2. Furthermore, I find the validation exercises presented (section 2.3) not very informative, and I feel that the case studies (section 4) are only useful to demonstrate model application, and do not yield significant scientific insight. Several of the conclusions offered based on the latter appear unfounded.

We appreciate the Reviewer's point of view, but the case studies were chosen precisely to "demonstrate model application"--this is Geoscientific Model Development after all, a journal founded in recognition that model description papers are significant and necessary science. That said, we have made a number of modifications to section 4 that we feel have put into better context both of the validation and case studies.

3. Finally, the model description is incomplete, as no details on implementation are given, short of the model code itself. This said, the model as such seems to be correct, and my comments only concern its presentation, applications, and interpretation. I leave it to the editor to decide whether the limited material presented here warrants a publication in gmd.

Because Hector, the SCM model used here, is described in a number of previous publications, we have expanded short summary of its structure and capabilities and provide citations for readers interested in more detail (see section 2.1). This allows the current manuscript to focus on HIRM and its applications. With respect to HIRM itself, a new paragraph (lines 110 - 114) describes its structure and implementation details.

Specific comments

1. p1/25 Earth system models of intermediate complexity should be mentioned here, especially as they are referred to later in the paper.

EMICs have been incorporated into the manuscript starting in line 26.

2. p2/31 "SCMs can be characterized as either process-based or idealized climate models." All models, especially SCMs are idealized - a better word would be "abstract" as opposed to "process-based". However, it would be better to speak of IRF-based models, as the authors use the term "idealized SCM" synonymously, although other types of non-process-based models may exist.

This is an interesting point. We adopted this terminology from Millar et al. 2017 to be consistent with the language used by other simple climate modeling groups. This being said we hope the changes made to the introduction and in lines 35-39 address these concerns.

3. p2/40 Another key nonlinearity that should be mentioned here is the chemistry of CO_2 uptake at the ocean surface.

Thanks for this suggestion, as it is an important nonlinearity. We have added it to line 50.

4. p2/44 Are all idealized SCMs based on IRFs? It would be better (and avoid repetition) to say "IRFs used in SCMs are defined as. . .".

This sentence has been removed.

5. p2/48 AR5 mentions several types of IRF-model, but the main model used to calculate GWP represents the relationship between emissions and CO_2, not temperature.

Thank you for pointing out the inaccuracy here. Due to changes in this section, this sentence has been removed.

6. p2/53 "Idealized SCMs may exhibit biased results, however, due to their lack of nonlinear dynamics." It is not in general true that idealized SCMs, meaning SCMs that are not (fully) process-based and use IRF functions, lack nonlinear dynamics. There are SCMs that apply IRFs only to the quasi-linear parts of the system, linking these with equations that capture essential nonlinearities (e.g. Joos and Bruno, 1996; Strassmann and Joos, 2018).

This was a miscommunication on our part; please see (lines 37-40) for the modified text which better situates HIRM among other nonlinear IRF based climate models. Thank you for sharing these references–they have been incorporated into the manuscript.

7. p2/56 Using the RF simulated by a model with nonlinearities as input hardly counts as "incorporating nonlinear dynamics".

We've tweaked the wording of the sentence so that now is reads: "In this manuscript we document and demonstrate a new highly idealized IRF-based framework".

8. p3/65 (whole paragraph) I find it rather misleading to call the use of an existing model to provide input a "framework", given that there does not seem to be any real coupling, i.e. exchange of information at intermediate timesteps. If this is the characteristic that distinguishes this "hybrid approach" from other IRF-based models, it does so in a negative way. There are SCMs that represent the climate response with IRFs and allow for coupling with a carbon cycle component at each timestep, for example, the BernSCM model (Strassmann and Joos, 2018). BernSCM combines IRF-based components describing linear systems with nonlinear parametrisations to capture the essential nonlinearities of the carbon cycle-climate system, and expresses the IRF-components as a system of ordinary differential equations to allow for efficient integration in coupled mode.

This paragraph has been edited to clarify how HIRM can be used with more sophisticated process-based models. However, we do feel that the term framework is sufficiently general to cover the application we have described here.

9. p3/69 "incorporate the nonlinear dynamics. . . if the majority of the nonlinear dynamics of SCMs occur between the emissions to radiative forcing calculation" -it would be more adequate here to say that the IRF-model, which really constitutes the contribution of the authors, does NOT incorporate any nonlinearities, because there aren't any

Noted and this sentence has been removed as part of edits to clarify the presentation in response to referee comments. We aimed for the revised text to communicate that the non-linear dynamics are, indeed, incorporated only through the use of forcing time series from the parent SCM.

10. p4/96 (whole paragraph) It is true that the carbon-cycle-climate feedback could be included in the IRF. However, the resulting model would still have strong limitations. It is likely that such a model, being liearized, would give accurate results only for a limited range of forcing scenarios or time scales.

We agree, and mention this in the paper (see lines 123-129).

11. p4/114 “The end of the IRF was extrapolated with an exponential decay function” Please mention the decay timescale.

More details about the exponential decay have been added: line (143-144) now reads “This IRF had a length of 300 years, in order to ensure the IRF was long enough to be convolved with the RF inputs; the end of the IRF was extrapolated with an exponential decay function to a length of 3000 with a decay constant of 0.20”.

12. p4/120 “underlying assumptions about where the majority of the nonlinearities occur are true” - This simply means the climate component of Hector is linear, which is to be expected of an SCM and could be inferred by looking at the design of that model.

Hector incorporates a diffusive model for ocean heat uptake which is, by definition, non-linear. However, as we have shown here, this nonlinearity is relatively weak compared to the non-linearities in emissions -> forcing calculations.

13. p4/122 I don’t see what chemical processes could affect the relationship between RF and temperature, at least in an SCM.

We have revised the text in lines (154-156) to clarify. (Chemical processes can impact relationships between emissions and forcing.)

14. p5/126 “For each RCP. . .” this sentence is not very informative and could be dropped.

We believe that this text does provide important information about the validation test set up, however we have modified the text so that information is more clearly portrayed.

15. p5/149 As mentioned above, this finding is not surprising; it merely characterizes the Hector SCM and holds no scientific information on the physical climate as such.

This sentence is not intended to convey any information on the physical climate system but is intended to show the extent to which the HIRM emulation matches the original hector result. As noted above, such a good match is not a foregone conclusion.

16. p5/155 “difference of 0.0%” There are no significant digits in this number

Noted, please see line 172 where we have updated this to 0%,

17. p6/156 As for the additivity of temperature changes, the lack of nonlinearities in the Hector climate component is not a scientifically relevant finding.

As noted above, this is not a foregone conclusion.

18. p6/164 “In this analysis, however, the black carbon (BC), organic carbon (OC), indirect SO2 effects (SO2i), and direct SO2 effects (SO2d) RF input time series were varied.” It is not correct to only vary these components. The uncertainty of other forcings should also be considered. The uncertainty of CO2 RF, for example, though small in relative terms, is important due to the dominant contribution of that component. Leaving out these uncertainties will result in an overconstrained temperature range.

We certainly agree with the comments of the reviewer in terms of scientific principles. We note, however, that the purpose of this section is to demonstrate how the tool could be used (as part of the GMD model documentation paper), not to conduct a rigorous uncertainty analysis.

This has been noted in the modified text.

19. p6/169 “sampled at intervals of 0.04 W/m2” It should perhaps be mentioned that this sampling does not produce a plausible probabilistic distribution of the results, since the RF uncertainties cannot be assumed to be uniformly distributed. Since the authors do not make a probabilistic interpretation, this is not a major issue, however

Correct we do not produce probabilistic results and have added text in 1997 to explicitly state that this case study does not produce probabilistic results.

20. p7/189 “This shows. . .” due to the overconstraining mentioned, this result is not valid in my opinion. Consequently the exercise described is only useful as an illustration of using the model framework, as stated in the following sentence.

This is, indeed, an application of the model as noted. However, we disagree that this conclusion is not valid. We note that most historical aerosol forcing estimation experiments don't separately examine the effects of different aerosol components (e.g., Forest 2018) - although Tomassini et al 2007 and Meinshausen et al. 2009 are exceptions and these citations have been added (although numerical values are not available making it difficult for us to compare directly). Most of these experiments generally scale total aerosol forcing up or down. In this exercise we have full forcing time series with different time paths for each forcing component (see Smith and Bond 2014 for examples). So we believe that our statement of what uncertainty components contribute to the constrained forcing range is valid as stated.

21. p7/193 (section 4.1) It is possible to use an IRF for a specific component from another model, as the authors do, but I am not sure how meaningful this is, since this mixes the climate responses of two different models. To get a consistent model emulation the IRFs for the other RF components should, in principle, be taken from the same model (i.e., NorESM-1).

We agree that a fully consistent emulation of any given climate model would entail consistent IRFs (and any other parameters) for all species - however there is little evidence

that large-scale climate responses are related to aerosol forcing responses in models (e.g., aerosol forcing is not correlated with climate sensitivity in CMIP6 models - Smith et al. 2020 - https://doi.org/10.5194/acp-20-9591-2020) so the experiment we present is reasonable as a sensitivity exercise.

Given that two coupled models so far have shown a dramatically different shape for the BC impulse response, our goal is to examine the impact of changing the BC IRF on global temperature. We, therefore, wish to use a realistic BC impulse for this calculation as derived from NorESM instead of postulating something more hypothetical.

22. p8/239 "it demonstrates nonlinear dynamics" I find this claim unfounded since the nonlinearities in question concern the dynamics of a previously existing model, while the model component contributed by the authors cannot represent the relevant nonlinearity, i.e., that of the climate-carbon cycle feedback.

Correct, HIRM depends on the parent model–in this case Hector. This limitation has been added to the manuscript in lines 307. We have clarified the wording.

23. p9/253 (whole paragraph) "most of the linearities" - there is no finding about any specific nonlinearities, and the fact that the Hector SCM has a linear RF-temperature response is no basis for a recommendation for further model development. The nearlinear relationship between RF and temperature is well known and has been demonstrated and exploited in SCMs for a long time (e.g., Joos and Bruno, 1996).

This paragraph has been modified to reflect this comment.

24. p9/271 "we demonstrate that the use of a forcing-based impulse response function overcomes most of these limitations." I don't see that any limitations are overcome by this approach.

Noted, this text has been removed.

25. p9/273 "These findings imply. . ." Again, there is no basis for such a recommendation.

Noted, this text has been revised so as to clarify the implications of this work.

Technical corrections

- p5/132 "In this experiment HIRM was configured" - The word "was" seems to be superfluous here.
Removed.

- In Table 1, the unit should be given.
Added.

References

- Joos and Bruno, 1996. Pulse response functions are cost-efficient tools to model the link between carbon emissions, atmospheric CO2 and global warming, Phys. Chem. Earth, 21, 471–476.

- Strassmann, K. M. and Joos, F.: The Bern Simple Climate Model (BernSCM) v1.0: an extensible and fully documented open-source re-implementation of the Bern reduced form model for global carbon cycle–climate simulations, Geosci. Model Dev., 11, 1887– 1908, https://doi.org/10.5194/gmd-11-1887-2018, 2018

Thank you for including these references, very helpful.