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

Anonymous Referee #3

The authors present a model that couples radiative forcing (potentially from any source, but here solely from the Hector simple climate model) to an impulse response function to calculate global mean surface temperature anomaly, with the possibility to choose different impulse response functions for different forcings (e.g. black carbon).

It is difficult to see what kind of advance this model is. HIRM relies very heavily on Hector, the details of which are documented extensively elsewhere. Using different IRFs for different forcings is not a new concept either (e.g. Richardson et al 2019 for CO2, CH4, BC, SO2 and solar forcing, Larson and Portmann 2016 for volcanic as a special case).

Indeed, the idea that there is a different climate response depending on the forcing agent has been around for quite some time. For example this has been incorporated mechanistically for aerosols in the MAGICC model for around 30 years now (Wigley. and Raper 1992), and more recently inferred by Shindell (2014) from GCM results. We have expanded our text to add this context. We note that this is a model description paper; we are not claiming that our results are new scientific advances, but rather are presenting this model as a useful tool for rapid analysis and as a testbed for model development and analysis. This is explicitly within scope of journals such as GMD.

This could be a nice module to include in Hector as an alternative way to calculate temperature in that model. The comparison to the default Hector temperature response function is shown in Figure 2 and seen to be almost identical, so this simplification (is it a simplification?) may be worthwhile. But, due to its nearly total dependence on Hector and the fact that species-dependent efficiacy response functions have been done previously, it doesn't qualify as a brand new model for me. Rather it is a submodule of Hector.

We have clarified in the revised text that HIRM is independent of Hector, it can be used with input from any model and with any IRF (lines 57-60). We use Hector because it is also open source (so the work here can be replicated by anyone who wishes to do so) and has a convenient interface for obtaining radiative forcing and temperature time series.

It is possible that there is more to this paper than meets the eye, but if there is it should stand out more, and if the authors believe this does warrant a standalone model, expend some effort in decoupling it from Hector and explain what is improved or new over e.g. Richardson et al. 2019

We hope the revised text does this. Note, however, that we do not claim that this work is improved over works such as Richardson et al. (2019), which presents an analysis of results from complex models (both coupled and atmosphere-only using a slab ocean). HIRM could certainly be used to quickly examine the implications of the IRFs derived in work such as

Richardson et al., but that work is of a different nature than what we present here in a model description paper.

Specific comments:

Line 18: Examining the effect of aerosol forcing on global temperature: a worthwhile cause. There is not actually done in this paper however. To me this would involve varying the present-day forcing of aerosols, climate sensitivity, and carbon cycle feedbacks, and investigating how this would cause temperature to evolve in a probabilistic fashion. The projections shown in figure 3 are far too constrained as explained in a later comment.

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.

Around line 25, there is a missing link between ESMs and SCMs: Earth System Models of Intermediate Complexity (EMICs). In fact, you could say that in decreasing order of complexity we have ESMs > GCMs > EMICs > SCMs.

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

In the paragraph starting on line 32, the authors discuss the differences between process-based and idealized simple climate models, presumably as a preface for introducing their own model that couples the two components. It is not clear to me that these two concepts are necessarily separate, and if they are, this model may not be as novel as the authors claim. The later versions of FAIR (Smith et al., 2018, GRL) include an impulse response function for CO2 emissions to concentrations and for converting forcing to temperature, and "processed-based" representations of concentrations of greenhouse gases, radiative forcing of GHGs and short-lived climate forcers, and feedbacks from temperature on the carbon cycle and radiative forcing. Leach et al. (2020) in the Generalised Impulse Response model extends the impulse-response framework of the carbon cycle to other greenhouse gases and short-lived climate forcers. In both of these models, the radiative forcing is internally calculated (process-based, in the language of the authors?) and not supplied externally/provided by a different model (as discussed in lines 82-83 for HIRM). For my benefit if not others, could you cite maybe one example of a "process-based" SCM on line 32 if these concepts are separate? MAGICC perhaps?

We have included MAGICC as an example of a process-based model in line 33. Furthermore the changes made to second and third paragraphs of the introduction section help clarifies the discussion regarding process based and impulse response function based models.

line 37: "top of troposphere" - this is an old and incomplete definition of radiative forcing, and effective radiative forcing is now preferred - the Smith et al (2018, JGR) reference which is in the bibliography (but not in the text, oddly) goes into some detail on this. I should say this discussion is of limited importance for SCMs.

We have expanded this discussion slightly to clarify how radiative forcing is defined in Hector (and, therefore, by inference in the examples given in this work).

line 54: "in addition, the physical interpretation of their behaviour is not always straightforward". Please explain why not.

This text was confusing and has been removed.

Line 90: This is a confusing paragraph. On first reading it seems like HIRM doesn't allow for species-dependent efficacy. Then I later read the discussion on BC, and see that the different IRF for BC can be included, which *is* a different efficacy (and time-dependent too). Then on second reading I see that the authors are talking about Hector not having species-dependent efficacy which is more evidence that this model is a component of Hector and not standalone. In general, in section 2.2 it is difficult to follow what the authors did. A flow diagram could help.

Yes, this section was unclear. HIRM can be configured with a unique IRF for each RF agent. However, for the purposes of the validation exercises HIRM had to be set up the same way as Hector, which only exhibits a single IRF. Text clarifying this point has been added in lines (86, 111, and 307).

Lines 97-102: It is not correct to exclude carbon cycle feedbacks. It is no good if HIRM can emulate Hector with feedbacks switched off if the latter is not representative of the real world. If the forcing comes from Hector in the first place, feedbacks need not be excluded in the Hector configuration, because HIRM doesn't calculate forcing. The analogy here would be concentration-driven GCM runs, where the concentrations are calculated by MAGICC (which includes carbon cycle feedbacks) but the GCMs themselves do not, going from concentrations > forcing > temperature.

We've modified the text from lines (123-129) where we discussed the decision to exclude the carbon cycle responses. For the purpose of our validation experiments it is appropriate to exclude them here, but they could be important to include in other applications.

line 120-121: people have underlying assumptions, but software doesn't

We respectfully disagree, but this is perhaps just an issue of semantics. Perhaps a better way of expressing it would be that programmers encode their assumptions into the software they create.

line 124: Hector's IRF - this is the function in figure 1 isn't it, because Hector is not impulse-response based. Could just refer to confirm.

Correct. Figure 1 is Hector's IRF that was obtained from by running hector as described in the text.

lines 131-132: it goes without saying that 4xCO2 tests the climate model's response to CO2. The importance of the 4xCO2 experiment is that it can be used to (imperfectly) estimate climate sensitivity, climate feedback and CO2 radiative forcing in GCMs (Gregory et al 2004) by using a forcing with a high enough signal-to-noise ratio to get a clear signal but still small enough to avoid substantial non-linearities and tipping points. Hence it can be used as a line of evidence for climate sensitivity, which itself is an input parameter to many simple climate models. Also, putting the Schwarber reference in line 132 reads like they invented this experiment.

Correct, it was not our attention to present the Schwarber reference like they invited this experiment, but we see how it could have been interpreted this way. We have modified the text to add more appropriate context.

line 145: The difference ... and the following sentence, can be dropped. It's apparent from the naked eye that the differences are imperceptible, I don't think this needs to be rigorous. Similar sentence in following paragraph.

Noted, text in this section has been changed.

line 157: Which SCM? Hector?

Noted and corrected.

lines 161-162: needs a reference

We now cite Forest 2018.

line 174: 29000 is a bit of a random number, is there a motivation for this?

29000 is the number of combinations of the uncertainty scalars when the uncertainty rangers were sampled, this is described in the text starting in line 203.

lines 176-177: wrong values (-1.9 to -0.1 is AR5 "very likely" i.e. 5-95% estimate), and also wrong citation (Myhre et al., 2013).

Correct, our range was modified from Myhre et al. 2013, the text in lines 212-213 more clearly describes the range that was used.

Figure 3: why does uncertainty reduce over time? 2100 temperature is very tightly constrained, but this is the timeframe over which uncertainties in radiative forcing, climate sensitivity and carbon cycle feedbacks multiply. I know these are not sampled, but this should very clearly be stated and the fact that this is not a true future uncertainty quantification of warming. I'd also check the constraints - is 1.6C of warming in 2010, which passes the constraint, realistic?

The uncertainty decreases by year 2100 because overall aerosol forcing decreases. In the scenario that was used aerosol and precursor emissions decrease substantially over the 20th

century so that, regardless of what is assumed about forcing per unit aerosol/precursor emissions, the overall impact of aerosol forcing is smaller. Therefore, the absolute magnitude of the impact of aerosol forcing uncertainty in 2100 also decreases.

Thank you for the careful reading of the figure. The upper constrained value of around 1.6C in 2010 is consistent with the applied temperature constraint which is only over 1880–2012. There is a non-trivial amount of positive forcing prior to 1880 due to both well-mixed greenhouse gases and also, potentially, from aerosols (If BC forcing is strong and cooling aerosol forcing is weak). See Smith and Bond 2014, Figure 4. The figure caption has been amended to note this.

lines 189-190: Important point, long known. Compare/cite some relevant studies e.g. Forest (2018). Figure 4 would be more naturally expressed in terms of a W/m2 aerosol forcing posterior for e.g year 2010, perhaps add a subpanel e. This would back up the claim that strong aerosol forcing values do not pass the constraints.

Most of these previous studies do not represent uncertainty in the different aerosol components separately. Only Meinshausen et al. 2009 and Tomassini et al 2007 include BC, OC, SO2 direct, and aerosol indirect forcings (Forest 2018, Table S4), but only show them graphically; we have added those references. Note in this example application the aerosol forcing range is supplied as a constraint so its not an independent output to compare to these previous results.

Figure 4 and its caption has been updated as per your suggestion.

Line 214: Mention the perturbation size from line 220 here. Richardson et al. (2019) included a multi-model response for BC and would be more appropriate to use than the single-model study here.

We agree this would be useful to use the Richardson et al response function. However, the parameters of their multi-model response function for BC was not provided in their paper or SI (and is not yet available from the authors). The perturbation size is mentioned in line 262.

line 230: maybe a better wording would be "... the global temperature was 0.2C lower using the specific BC IRF from Sand et al. (2015)." Avoid using "decreased" in this sentence.

Noted and changed.

minor:

line 9: Earth (rather than earth)
line 29-30: would probably get picked up in proofing but check citation spellings (Meehl, Meinshausen)
line 40: Myhre et al.

line 91: forging
Noted and changed.

line 200: also Richardson et al. 2019
Added.

line 226: units, should be C/(W/m2)? i.e. the m2 is in the numerator

It now reads °C $W^{-1}m^{-2}$

line 253: HRIM (and in several other places)
line 244: "a" not required

Noted and changed.

line 245: "dynamics" - not really dynamics is it - just a different IRF

Noted, this text changed to "temperature response".

line 438: 29000 times

Noted and changed.

References:

Richardson et al. 2019: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019JD030581
Larson and Portmann 2016:
https://journals.ametsoc.org/jcli/article/29/4/1497/35504/ATemporal-Kernel-Method-to-Compute-Effective
Smith et al. 2018, GRL: https://gmd.copernicus.org/articles/11/2273/2018/
Leach et al. 2020: https://gmd.copernicus.org/preprints/gmd-2019-379/
Forster et al. 2016: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016JD025320
Smith et al. 2018, JGR: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018GL079826
Gregory et al. 2004: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2003GL018747
Forest 2018: https://link.springer.com/article/10.1007/s40641-018-0085-2

Thank you for the suggested references.