Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-264-RC2, 2020


Geoscientific Model Development Discussions

# *Interactive comment on* "A computationally efficient model for probabilistic local warming projections constrained by history matching and pattern scaling" *by* Philip Goodwin et al.

**Anonymous Referee #2**

Received and published: 10 January 2020

## 1 General comments

Goodwin et al. present a tool for projecting local warming with uncertainty from multiple anthropogenic emissions scenarios. The major advance of the paper is the combination of output from a probabilistic climate model and warming ratios from AOGCM/ESMs (I note that the MAGICC/SCENGEN, http://www.cgd.ucar.edu/cas/wigley/magicc/, tool does a similar thing but given that this paper is not tightly coupled to MAGICC or any other probabilistic climate model and its code is open sourced I consider this paper to be a significant advance on the MAGICC/SCENGEN tool). I feel that this advance could be a very useful addition to the



literature if a few concerns were addressed to provide more confidence in the paper's conclusions.

My major concerns focus on: whether the tool is actually scenario specific or not, how uncertainties from the climate model and LGRTC are combined and and whether WASP is actually a key part of the tool or whether any probabilistic climate model could be used.

One other key comment, given the availability of CMIP6 model output, I feel this paper could be significantly improved if it were to use CMIP6 output rather than focussing on CMIP5.

## 2 Major concerns

### 2.1 Scenario specificity of pattern scaling

It is not clear to me that the pattern scaling technique here is actually scenario agnostic. All the presented results are scenario specific (the RCP45 projections use RCP45 LGRTC and the RCP85 projections use RCP85 LGRTC) and there is no analysis of whether a 'general LGRTC' can be used nor whether such a 'general LGRTC' would have small enough uncertainties as to be useful.

I feel the comment (page 6, line 10), 'This allows future users to choose the spatial pattern scaling that is most suitable for their scenario.' is misleading. Only 3 patterns are available and none of them have been shown to be applicable for an emissions scenario different to the one from which they were derived (see comment above). Such cross-validation would be a vital step to providing confidence that the spatial pattern derived from one scenario can then be applied to any arbitrary scenario.

I am not convinced by the comment (page 4, line 8), 'The absolute value of differences

in LGRTC between the three scenarios was below 0.72C per°C in all grid-cells and mostly below 0.2C perC over the continents. Therefore, the choice of the emission scenario to define spatial pattern of warming in this study is not much relevant when only inhabited regions are considered.' Relative to strong mitigation targets (e.g. the 1.5°C target), I am not convinced these are trivial variations. In addition, in this context 'mostly' is meaningless and provides no quantification of how wide the disagreement is nor of the regions in which this generalisation doesn't hold (and how wrong it is).

I am also not convinced by the comment (page 4, line 19), 'This might have led to the large differences in the Arctic region, but detailed analysis and explanation is outside the scope of this study.' If the pattern scaling approach is to be used for arbitrary scenarios, there needs to be evidence that one pattern, with sufficiently large uncertainties, can be applied to multiple scenarios and give results that are in line with known results from CMIP models. Any differences need to be explained as they are of key interest when applying this tool (or the tools' domain of applicability should only be limited to those regions where the differences are small/well understood).

I think the data is there to address this concern. One suggestion (which would satisfy me) would be to derive some 'general LGRTC' (including uncertainty) which could be used with any emissions scenario. The 'general LGRTC' could then be applied to the RCPs (here meaning all RCPs, including RCP26 and RCP60, not just RCP45 and RCP85) and the differences quantified. This would provide a meaningful quantification of how big the uncertainties need to be on a 'general LGRTC' for it to sufficiently capture the variation across CMIP models and scenarios in the cases where we have data. I would be even more convinced if a 'general LGRTC' derived from CMIP5 RCPs was shown to hold for CMIP6 SSP scenarios.



2.2 Scenario specificity of WASP

WASP currently requires exogenous estimates of non-CO2 radiative forcing (see manuscript paragraph starting page 7, line 33). As far as I can tell, this means that this tool is not applicable to arbitray emissions scenarios but rather only ones for which there is an available non-CO2 radiative forcing quantification. I feel this is a rather fatal flaw of a tool which is meant to be applicable to arbitrary emissions scenarios.

An easy remedy would be to alter the tool from being 'WASP/LGRTC' to 'a general framework for coupling probabilistic climate model output and LGRTC' (insert acronym here) i.e. remove the explicit dependence on WASP. I can't see any reason why WASP is the only model with which this tool would work. This paper could still illustrate the use of the framework with WASP output, but such a reframing would make clear that the coupling could be done with any probabilistic climate model so a model which can run fully GHG-emissions driven could be used instead and would immediately fix the issue of WASP's limited available scenario set.

2.3 Combination of uncertainties

I am not convinced that the combination of uncertainties in equation 2 is correct. In equation 2, shouldn't the resulting distribution be the product/convolution of the two distributions rather than the output of random sampling from the two distributions? Given LGRTC is assumed to be gaussian, and that the WASP output is approximately gaussian, wouldn't it be better to derive the distribution of Delta T_i (x, y, t) by taking the product of two gaussians (see e.g. https://ccrma.stanford.edu/~jos/sasp/Product_Two_Gaussian_PDFs.html) which isn't the same as the product of two gaussian variables (see e.g. https://math.stackexchange.com/questions/101062/is-the-product-of-two-gaussian-random-variables-also-a-gaussian).

I'm happy to be corrected on this as I am not a statistical expert. However, regardless of whether I am correct or not I think some explanation must be added to the manuscript or the supplementary to explain this uncertainty propogation.

2.4 Reliance on WASP

It is not clear if this paper is using an existing WASP probabilistic distribution or presenting a new one (e.g. contradiction between page 5, line 9: '3x$10^6$ members' and page 2, line 23: '$10^8$ simulations'). If the reframing suggested earlier were to take place then this is no longer an issue (as the choice of particular probabilistic climate model is just for illustration and isn't a key feature of the tool). However, if this particular WASP probabilistic distribution is key then I would have to consider that component more closely.

(If the WASP probabilistic distribution is not key this entire paragraph can be ignored but for completeness) At the moment my only question is about the Monte Carlo sampling. Supplementary Table 2 of Goodwin et al. 2018b shows 18 parameters. With 3x$10^6$ members you're effectively taking a bit over 5 steps in each parameter axis ($18^5$ ~ 2x$10^6$). This appears to be a fairly sparse sampling, which could be a problem no? I wasn't convinced by Goodwin et al. 2018b, 'This observation-consistent ensemble displays good agreement with the full ranges for all the observational quantities (Supplementary Table 4), which demonstrates that the 3x$10^4$ simulations have a good coverage of observational parameter space.' It seems perfectly plausible to me that the 95% ranges could agree but the distributions themselves are otherwise very different. If you've considered this before and can include the answers in the paper or point to them in the paper that would be great, if not then a sentence highlighting this and saying that they're areas for future research would suffice.



## 3 Specific comments

1. 'Thus a TCRE framework is applicable for certain situations, including idealised scenarios where the TRCE has already been established, but in the general case a time-dependent Earth system model is required.' (page 2, line 13) Can you make some comment about what this means for the TCRE framework that was heavily used in SR1.5 (see Rogelj et al. Nature 2019, 'Estimating and tracking the remaining carbon budget for stringent climate targets')? For example, does it mean that the framework can only be applied if its components were derived with a suitable scenario set?

2. Opening paragraph, the commas mean that the sentence says 'The dominant climate projections are made using the Climate Model Inter-comparison Project phase 5 ensemble'. Given CMIP6 is well and truly underway, can you re-write the entire paragraph to make CMIP references more general e.g. 'The dominant climate projections are made using results from multiple phases of the Coupled Model Inter-comparison Project' (the second sentence of the paragraph also needs a similar adjustment)? (See also comment about using CMIP6 data throughout the manuscript)

3. 'ideal tool for future incorporation into an Integrated Assessment Model framework' (page 1, line 25). Given WASP's requirement of exogenous non-CO2 forcing I don't think this is true so would remove this phrase. It could, of course, remain if the reframing towards a more general probabilistic climate model-LGRTC framework suggested above was made.

4. The paragraph beginning page 4, line 11. I was very confused by the entire discussion of reference periods and comparison periods throughout this section. Specifically, 'we have chosen the preindustrial climate as the reference period in 1pctCO2'. What does this mean? Is the reference period in 1pctCO2 1850-1900

or something else? ‘in the RCP scenarios we used beginning of the 21st century’. Is my understanding right that you are saying you used the beginning of the 21st century as the pre-industrial climate for the RCPs? If yes, this seems a very odd choice. ‘end period... years 2079-2098 in the RCPs’, this seems a very odd choice of end period, why not 2081-2100 as is used in the IPCC reports?

5. why are the TCRE fits a) quadratic and b) only done to one scenario? Given the use of the TCRE concept throughout the literature, either a linear, scenario-independent fit between warming and cumulative carbon emissions should be done or a much more thorough discussion of why a scenario-dependent, quadratic fit is appropriate should be added.

## 4 Figures

Figure 1 caption: is this the multi-model mean?

Figure 2: a different colour palette and increments would be helpful so you can see whether the standard deviation is in 0-0.3 or 0.3-0.6, very hard to tell at the moment and such a difference are of interest with respect to the renewed focus on 0.5C temperature increments following SR1.5

Figure 4: add standard deviation panel too please so the size of the uncertainties is immediately obvious (doing the differences by eye is basically impossible given how wide the colour bar scale is)

## 5 Technical corrections

page 1, line 24: delete ‘arbitrary’



page 1, line 27: ‘tool making’ → ‘tool for making’

page 1, line 31: ‘CMIP’ stands for ‘Coupled Model Intercomparison Project’, not ‘Climate Model Intercomparison Project’

page 1, line 34: ‘carbon-emissions’ → ‘anthropogenic-emission’

page 2, line 2: ‘resolutionand/ora’ → ‘resolution and/or a’

page 2, line 6: ‘and an’ → ‘and a’

page 2, line 6: ‘hemispherical averaged’ → ‘hemispherical land-ocean averaged’

page 2, line 9: ‘is applied’ → ‘can be applied’

page 2, line 10: do you have a reference for SCENGEN? Perhaps http://www.cgd.ucar.edu/cas/wigley/magicc/?

page 2, line 12: ‘IMAGE, MESSAGEframeworks’ → ‘IMAGE and MESSAGE frameworks’

page 3, line 11-12: ‘for example where cooling following negative emissions may not re-tracethe previous warming pathway’ needs to be re-written, I can guess what you’re saying but the sentence doesn’t actually make sense

page 4, line 5: delete ‘such’

page 4, line 11: ‘Despite that’ → ‘Despite the fact that’ ? I’m not actually sure what is intended here.

page 4, line 22: ‘RCP scenarios more usable than those from the 1pctCO2 scenario to be used to predict warming patterns in the 21st century’ → ‘RCP scenarios more appropriate those from the 1pctCO2 scenario to predict warming patterns in the 21st century’

page 4, line 27: Missing units on the numbers

page 4, line 30: This paragraph is a very roundabout way of saying that the projections from each CMIP model are internally physically consistent, but as the WASP/LGRTC uses averages of each CMIP model, its results are unlikely to be internally physically consistent. Can you cut the paragraph to one or two sentences?

page 4, line 39: Delete this paragraph

page 5, line 17: ', (Figure 3)' → ' (Figure 3)'

page 6, line 15: 'function' → 'a function'

page 7, line 17: Calling WASP an 'efficient ESM' is a stretch, use 'efficient climate model' or 'simple ESM' or 'ESM emulator' instead

page 7, line 23: Straw-manning paragraph, the WASP part is not new (lots of other emulators with probabilistic distributions constrained to history), only the LGRTC scaling is

page 7, line 38: Given lack of any tests for other platforms or docs for how to set things up, I am dubious of the phrase, 'Both the WASP/LGRTC model and the quadratic approximation to WASP/LGRTC model are easy to use', particularly in relation to WASP's useability. Given all the plotting is all MATLAB based you have a significant barrier to entry.

### 5.1 Supplementary information

page 3, Table caption: Am I correct in guessing that the '90 to 95%' confidence interval means that 95% confidence interval was used where data was available otherwise 90% confidence interval was used? If not, can you please clarify what a range of confidence interval means as I haven't seen such a usage before (I would have thought it would just be '95% confidence interval')?