Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-264-AC1, 2020

Geoscientific Model Development Discussions
Open Access

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

Philip Goodwin et al.

p.a.goodwin@soton.ac.uk

Received and published: 12 February 2020

Reply to SC1: Authors' response to the editorial comment:

The comment by Astrid Kerkweg drew attention to an editorial requirement:

Reviewer's comment: *"In particular, please note that for your paper, the following requirement has not been met in the Discussions paper:*

"The main paper must give the model name and version number (or other unique identifier) in the title."

Please add the names/acronyms (WASP/LGRTC) of the models used/developed and

Printer-friendly version

Discussion paper

their version numbers to the title upon your revised submission to GMD. Yours,

Astrid Kerkweg"

Authors' response: Thank you for drawing this editorial requirement to our attention. When revising the manuscript, we will include in the manuscript title a unique model name and version number for the combined WASP/LGRTC model for spatial warming patterns.

Reply to RC1: Authors' responses to reviewer 1's comments:

We thank reviewer 1, Dr. Christopher Smith, for important and insightful comments about our manuscript. Below we explain how we will use these comments to further improve our manuscript.

Reviewer's comment: *"General comments This paper describes a simple methodology for translating global mean surface temperature diagnostic output from a simple climate model (WASP, but in theory any model like MAGICC, FAIR, Hector could theoretically be used) into regional surface temperature changes using a pattern scaling approach. While this is not a necessarily new concept (see fldgen: https://www.geosci-model-dev.net/12/1477/2019/), it is appreciated that a quick and simple tool would be greatly useful for translating the output of simple climate models (e.g. from IAMs) to regional impacts. Additionally, there is a nice link from carbon emissions/carbon budgets to future carbon emissions. With this knowledge it could be possible to assess regional impacts as a function of the remaining carbon budget (e.g. to 1.5C)."*

Authors' response: Point 1: Applicability of the spatial tool to any model capable of projecting global mean warming. We agree that any model capable of generating probabilistic projections for global mean surface temperature could be combined with the spatial tools presented in this manuscript. This would then generate projections for the mean and standard deviation for future local warming. In a revised manuscript we will highlight how the spatial tool can be coupled to any model projecting global mean

Printer-friendly version

Discussion paper

warming, and not just the WASP model. This change will make the manuscript more general in nature and of interest to a wider range of readers.

Point 2: Link to emissions budgets and regional impacts. We agree that the approximation tool presented provides a useful link to assess regional impacts from carbon emissions/carbon budgets. In a revised manuscript we will improve this link by exploring the spatial warming pattern (the LGRTC) for the RCP2.6 scenario – a scenario with strong mitigation and a high likelihood of meeting the Paris Climate Agreement goals.

Reviewer's comment: *"p4 l4-10: I am not sure if three scenarios that all show various rates of continually increasing warming are sufficient to make this conclusion. I would suspect that this does not hold for RCP2.6 where most models stabilise in temperature but regional patterns may continue to evolve. It would be good to show this. It would be helpful to see the 1pctCO2 scenarios for comparison in figure 1, also. (also relevant to p6 l6-10)"*

Authors' response: Scenarios with increasing warming and scenarios with stabilised warming. We agree that the scenarios considered have increasing warming although we note that in RCP4.5 the there is little additional warming after 2080 across the 13 CMIP5 models considered by Goodwin et al., (2018b – see figure 2c therein, grey shaded area), we agree that there is little time for the warming pattern to continue to evolve.

We also agreed that the RCP2.6 scenario offers a chance to explore our LGRTC tool for generating local warming projections in a scenario with stabilised climate. In a revised version we will produce a LGRTC for RCP2.6 and compare this to the existing scenarios. This will allow us to test our methodology for a climate with warming stabilised close to the Paris Climate Agreement warming targets of 1.5 and 2.0 °C.

Reviewer's comment: *"p4l22: a point on different non-CO2 forcers in the three scenarios - the RCPs are quite heterogeneous in their aerosol forcing in future scenarios,* *and 1pctCO2 does not include them. I'm not sure this gives us much information for pattern scaling for custom emissions scenarios. See figure 3 in Liu et al. for temperature responses to - admittedly somewhat extreme - cases of aerosol forcing in Europe and Asia. https://doi.org/10.1175/JCLI-D-17-0439.s1 . Some more discussion about how this model could handle widely varying timeseries of global and regional aerosol forcing would really help strengthen the model (and paper)."*

Authors' response: Pattern scaling for custom emissions scenarios with extreme aerosol emissions. In the production of the RCP scenarios, assumptions have been made as to the relative amounts of different forcing agents emitted (e.g. CO2, other well mixed greenhouse gasses, aerosols etc) (e.g. see IPCC, 2013 or Meinshausen et al., 2011). Some forcing agents, specifically aerosols, are not well mixed in the atmosphere but exert a significant local influence. We agree that for extreme cases of localised aerosol radiative forcing the pattern of warming would be different to the assumptions used in the generation of the different RCP scenarios.

In a revised manuscript, we will discuss how this approach could be extended in future study to account for custom scenarios containing extreme aerosols forcing. The work itself is reserved for future study because, fundamentally, one needs additional information from complex climate models involving separate model integrations with non RCP (or SSP) style scenarios. Briefly, a revised manuscript will discuss how the method presented could be extended in future work by: (1) Calculating the LGRTC for a scenario with CO2 only (or well mixed greenhouse gas only) forcing, (2) Exploring the sensitivity of LGRTC patterns to regional aerosol emission in complex climate models, for cases with idealised aerosol emissions for each region (e.g. Europe, Asia etc). (3) Combining the LGRTC from well mixed greenhouse forcing with the LGRTC from idealised aerosol forcing in each region, using knowledge of the relative emissions from greenhouse gasses and aerosols by region in the custom scenario, to generate a custom warming pattern for a scenario with extreme aerosol emission patterns. The key to such a method working is that the uncertainties introduced by the assumptions involved in combining different spatial patterns are smaller than the uncertainties introduced by the range of CMIP class model responses for a given scenario.

Since this method relies on new targeted experiments with complex CMIP models, this is not feasible to conduct for this paper. However, we will spell out the possible methods by which this could be achieved in future work in the revised version.

Note that the calculations for the LGRTC presented here will apply for custom scenarios that do utilise similar assumptions to the RCP scenarios in terms of the relative amounts of well mixed greenhouse emissions and aerosols.

Minor comments:

Reviewer comment: *"Minor p1l36: a couple of years: I'd say it's more like 7 or 8, approximately in line with the corresponding IPCC Assessments."*

Authors' response: Agreed, we a revised manuscript will be amended accordingly.

Reviewer's comment: *"p2l25: I think it's worth explaining which observational temperature datasets were used because the future projections you obtain from history matching will depend on whether they are observational blended global near-surface air temperautre and sea surface temperature and whether they infill for missing data."*

Authors' response: Agreed – the global mean temperature datasets used are blended land air temperature and sea surface temperature records (e.g. HadCRUT4 and GISTEMP). We will spell this out in a revised manuscript so that readers understand precisely what our global mean temperature anomaly projections should be compared to.

Typographic/stylistic points. We also thank Dr. Smith for also raising minor/stylistic points. We confirm that we agree with all minor/stylistic comments raised by reviewer 1 and can confirm that these will all be addressed in a revised version.

Reply to RC2 (as updated in EC1): Reply to reviewer 2's comments

Reviewer comment: *"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."*

Authors' response: General comments. We are pleased the reviewer sees the advance offered, and in a revision we will amend the manuscript to address the specific concerns – please see details below.

Reviewer comment: *"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."

Authors' response: Concerns. We spell out in detail below how we will update a revised manuscript to address the concerns raised. In brief, in a revised manuscript we will: (1) Analyse the LGRTC for an additional scenario, RCP2.6, and provide more robust statistical comparisons of the differences in LGRTC for the different scenarios, including identifying a spatial domain over which a single LGRTC can be applied; (2) Stress that the method is not specific to the WASP model, and make it clear that our methodology can be applied to any efficient model generating projections of global mean surface warming; and (3) We will reserve the analysis of CMIP6 model output for

Printer-friendly version

Discussion paper

future study.

Reviewer comment: *"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."

Authors' response: Scenario specificity of the LGRTC. Agreed that an analysis of whether a 'general LGRTC' can be defined will improve the manuscript. We will include such an analysis in the revised version (see below for details)

Reviewer comment: *"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."*

Authors' response: Agreed that the sentence "This allows future users to choose the spatial pattern scaling that is most suitable for their scenario' is unclear in its present form. This statement will be removed in a revised version. See below for details on broader points.

Reviewer comment: *"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 perC 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 1.5C target), I am not convinced these are trivial variations. In addition,*

Printer-friendly version

Discussion paper

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)."

Authors' response: Imprecise wording of comparisons between scenario-LGRTC patters. Agreed that the highlighted sentences do not provide robust statistical analysis of the differences and similarities between the LGRTC patterns for the scenarios. In a revised manuscript, we will provide a robust statistical comparison of the LGRTC for different scenarios over different areas of the domain. We will define the domain over which the tool is applicable. See answer to the next paragraph for more details.

Reviewer comment: *"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."*

Authors response: Addressing concern over applicability of LGRTC for different scenarios/scenario dependence quantification. Agreed that greater insight into the amount of scenario-dependence is required, and that this affects the validity of using the LGRTC for scenarios other than the scenarios from which they were derived. In a revised manuscript we will:

(1) Analyse the LGRTC for an additional scenario, RCP2.6 (a stabilisation scenario with strong mitigation in line with the Paris Climate Agreement's targets of keeping warming under 2.0 °C).

(2) Provide a more meaningful statistical comparison of the LGRTC for the different scenarios (including RCP8.5, RCP4.5, RCP2.6 and 1 per cent CO2), including comparing the magnitude of uncertainty due to differences 'within scenario but between CMIP5 models' to the differences 'between scenarios'. i.e. comparing sigma_LGRTC within a scenario to the differences between mu_LGRTC for different scenarios.

This comparison will avoid language such as 'mostly' and be quantitative as to the differences between scenarios over various spatial domains.

(3) Explore the feasibility of defining a domain over which a general LGRTC can be defined (with uncertainties large enough to capture variation across CMIP5 models and variation between scenarios).

Ultimately, a key property of a single LGRTC must be that the uncertainty introduced by the scenario choice is less than the uncertainty introduced by the range of CMIP-class model responses within each a given scenario. In a revised manuscript, we will identify domains over a single LGRTC can be identified for all scenarios.

We will reserve comparisons to CMIP6 for future study.

Reviewer's comment: *"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."

Authors' response: The method's (non)reliance on WASP. We agree that the LGRTC tool can be applied to any arbitrary probabilistic climate model ensemble, not just the WASP ensemble used in the study. We will reframe the manuscript in terms of offering a general framework, with WASP the efficient model used to illustrate the tool.

Reviewer's comment: *"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-twogaussian-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."*

Authors' response: Combination of uncertainties. We agree with the statistical points made about the random sampling of two Gaussian distributions not in general giving the same answer as the convolution of two Gaussian distributions.

However, to make our LGRTC method applicable to any arbitrary probabilistic projection of global mean surface warming (not just from this WASP ensemble), we cannot assume that the projection of global mean surface warming is Gaussian. Therefore, to ensure that our approach is generally applicable to any efficient model's projection of global mean surface warming, we cannot take product of two Gaussian distributions as suggested by the reviewer.

Our method, of randomly sampling from both the distributions of global mean warming and LGRTC, is applicable to any arbitrary projection of global mean surface warming.

We also point out that in the MATLAB approximation tool, which does tie in to the WASP ensemble, we have used the product of two Gaussian distributions rather than random sampling.

In revision, we will clarify why we cannot assume a Gaussian distribution for future global mean surface warming if our approach local warming approach is to be applied to any arbitrary model for generating probabilistic projections of global mean warming. We will also state that it is an option to convolute the Gaussian distributions, in the special case where one already knows that the global mean warming distribution is Gaussian.

Reviewer's comment: *"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: '3x10^6 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."*

Authors' response: The (non)reliance on WASP. The novel methodology (of combining the LGRTC with a probabilistic ensemble of global mean warming from an efficient numerical model) is not tied to WASP. Therefore, we will be making the reframing suggested by the reviewer earlier clear in a revised manuscript. We adopt the probabilistic ensemble generated in Goodwin et al (2018b). We will make the particular ensemble used clear in the revised manuscript.

We note that there is not a contradiction between the 3×10^6 members of the posterior ensemble and 10^8 members of the prior ensemble in the WASP methodology (see below).

Reviewer's comment: *"(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 3×10^6 members you're effectively taking a bit over 5 steps in each parameter axis (18^5 2×10^6). 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 3×10^4 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."*

Authors' response: For completeness (with no bearing on the manuscript as method is not tied to the WASP model). There is no contradiction in the manuscript.

The generation of the probabilistic ensemble involves an initial ensemble of 10^8 simulations based on input parameter distributions that reflect prior knowledge of Earth system properties (such as climate sensitivity). These input parameter distributions are independent of one another (i.e. we do not assume that we know how one input parameter affects the value of other input parameters). From this initial ensemble of 10^8 simulations, $3x10^6$ are extracted to form the final probabilistic ensemble using an observational consistency test.

Some Monte Carlo sampling methods (effectively) ascribe a graduated weighting to each simulation according to the simulated position on the probability distribution of each of the observables used to constrain the system. For example if each simulated value agreed with the best estimate of each observable then that simulation would have the highest weighting. Here, instead of a graduated weighting, the weighting for each ensemble member is either 1 (included as observation-consistent) or 0 (excluded as inconsistent), based on whether the simulation agrees with the observable quantities to 95% confidence (see Goodwin et al, 2018b for a full methodological description including how the tails of the distribution are included).

Primarily, this accept/reject approach is taken because we assume we are able to ascribe some confidence range for each historical observational reconstruction, but we do not assume to know the shape of the probability distribution. To see why, consider the historic global ocean heat content anomaly. The observational consistency test in WASP for historic whole ocean heat content anomaly is based on 6 published reconstructions (e.g. see Goodwin et al., 2018b – see fig. 1 panels c and d therein for the six different reconstructions of ocean heat content anomaly considered). These reconstructions range from a best estimate of the increase in whole-ocean heat content from years 1971 to 2010 from under +200 ZJ to over +300 ZJ.

Considering this variation between the six reconstructions, a significant component of uncertainty in historic global ocean heat content anomaly is due to systematic uncertainty in the methodology used to convert sparse observations into a global mean, and not random uncertainty in the underlying sparse observations themselves. Producing a probability distribution for ocean heat content anomaly would thus entail a highly subjective judgement on the relative merits of each different methodology used by the six different observational reconstructions. To avoid this highly subjective judgement (required for a graduated weighting approach), we use an accept/reject approach, based on whether a simulation lies between the minimum to maximum values of the 95% uncertainty ranges for each of the six reconstructions. (i.e. a simulation is ‘observation-consistent’ with whole ocean heat content anomaly reconstructions if it lies within the 95% uncertainty ranges of any of the six reconstructions). Therefore, the 95% probability range is known, but the shape of the probability distribution is unknown.

Also, in terms of numbers of simulations in the ensemble it should be noted we have generated 10^8 simulations, from which we have extracted 3×10^6 simulations. Therefore, our initial sampling is adequate.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-264, 2019.