Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2017-36-RC1, 2017

Geoscientific Model Development Discussions

Interactive comment on "A method to encapsulate model structural uncertainty in ensemble projections of future climate" *by* Jared Lewis et al.

Anonymous Referee #1

Received and published: 8 May 2017

In this paper the authors have developed a method to encapsulate structural uncertainty in ensemble projections of future climate by combining regional climate model output with that from a simple climate model. The aim is produce a large ensemble of climate variables representing that which would be produced from a global climate model if it weren't made impossible by computational demands. A small ensemble of climate variables (T_min and T_max) and annual mean global temperature is produced from the regional climate model. The relationships between the climate variables and temperature are used to produce a large ensemble of the climate variables from the more readily available annual mean global mean temperature (simulated in the simple climate model). Observations are used to provide climatology, weather variability and maintain spatial coherence to the predictions. I think the underlying method that has been developed is interesting and would benefit the scientific community if published,

my concern about publishing in GMD is that this paper is about a method rather than a model. If the method has been developed such that it is a model in itself then the paper needs to be focussed on this and it requires more discussion about how users will run the model beyond what is presented in the paper, particularly with the new climate variables that will be included, how a new region is considered, how a new RCP is used and what will happen to the model when developments to the simple climate model and global/regional climate model are made. It may perhaps be better placed in a journal such as 'Advances in Statistical Climatology, Meteorology and Oceanography (ASCMO)'.

I will continue the review based on the method development.

General comment:

As a method I think this is a neat way to produce a large ensemble of climate variables at a regional level that would otherwise be unavailable and attempting to encapsulate the structural uncertainty and weather variability. In some places I find the method hard to follow and the paper needs clarity. In some places I think it would benefit from further equations to back up the text. I would also like to see more discussion on how parametric uncertainty could be included or why it is not deemed necessary. I would also like to see how additional RCPs are included in the method – are they included in the final PDFs or do you envisage separate PDFs for each scenario?

Particular points:

1. Introduction:

There is not enough of a literature review here. In particular, the UK Met Office have developed methods to produce probabilistic climate projections with a similar aim to this paper. Please discuss how this work achieves the goal differently. It would be good to see the work of Tebaldi and Knutti referenced here too discussing the difficulties in producing probabilistic information from multiple models.

In particular, the following papers should be referenced and discussed. Murphy, J. M. et al. A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. Phil. Trans. R. Soc. A 365, 1993–2028 (2007). Sexton, D., Murphy, J., Collins, M. & Webb, M.Multivariate predictions using imperfect climate models: Part 1 outline of methodology. Clim. Dynam. http://dx.doi.org/10.1007/s00382-011-1208-9 (2011). Harris, G. R., et al. "Probabilistic projections for 21st century European climate." Natural Hazards and Earth System Sciences 10.9 (2010): 2009-2020.

Line 2: 'will correctly simulate that trajectory'. I think the inclusion of 'correctly' is necessary since all climate models are attempting to simulate this trajectory.

2.1 Regional Climate Model

How dependent on the RCM are the relationships that are established? Would the relationships be expected to change with a new RCM?

Page 3, Line 10: What does 'adequate' mean? Page 3, Line 21: 'The model simulates all atmospheric and land surface processes'. There are missing processes and those that are included are subject to uncertainties – is the method robust to this? Page 3, Line 26: What is the implication of the remaining bias?

2.2 Simple climate model

Is MAGICC the only such model? How is it known to simulate annual mean global mean temperature adequately? Are the 19 AOGCMs and 10 carbon cycle models defined by MAGICC or could a user change what is included?

Page 4, Line 2: It is you that considers the 190 simulations as equally probable. 'We consider the resultant 190 different 'tunings' for MAGICC to be equally probable.' Could a user make different decisions? Page 4, Line 3: What does 'some' mean? Since the title of the paper states that the structural uncertainty is encapsulated this needs explaining. Page 4, Line 7: This is the first mention of 'local' - I think this should be

included earlier as it's an important point. Earlier, X is used. It would add clarity to define X better earlier on and use X here and in future.

2.3 Virtual Climate Station Network

Is this the best observational dataset beyond New Zealand?

3. Methodology

This section would benefit from some equations to clarify the text – especially right at the beginning to go with the bullet points.

Why is the period 1960-2100 used?

Page 4, Line 27: 'one or more RCM simulations' – I don't quite understand how the RCM simulations are being used here? Later it says you use five – what is the benefit of five and how did you choose these five? How could a user choose a different number? Page 4, Line 28: You haven't defined alpha yet so this is confusing. Can you state what the sampling of these alpha values is representing instead? Page 5, Line 2: 'valid for any GHG emissions scenarios' – is it valid if the regression isn't robust? How can a user validate this if they change the GHG scenario? Page 5, Line 2: Why is the anomaly calculated to the 2000-2010 baseline? You later say it is rather short so please say why you have chosen it.

3.1 Climatology

Page 5, Line 9: Are you following the method cited or have you expanded it? Can you say why you have chosen this particular method to account for seasonality? Figure 2: I think magenta and red are mixed up.

3.2.1 Training phase

I find this section particularly hard to follow and think it would benefit from more precise equations. Equation 2 doesn't help to explain what you actually did in terms of using time series from 5 RCMs. How have you built a model to explain the relationship be-

tween variables of a different scale (daily and annual)? How robust is this relationship with only 5 RCMs? Can you explain the statistical models that you build and how they are validated to produce robust relationships that can be used with the simple climate model output?

Page 7, Line 6: The relationship itself is calculated over the period 1960-2100? Page 7, Line 7: 'an annual mean global mean ..temp.. series' – do these match the period of X'? Are there five series from the five models? Page 7, Line 7: α(d) is the fit coefficient – are there 365 of these values? Page 7, Line 13: GCM 'and' RCM. Page 7, Line 17: I don't see how the Fourier pairs are embedded in equation 2 to find five fit coefficients from 365? Page 7, Line 18: 'recall, that this is applied at..' – I don't see where it told me that is is applied for each RCM simulation – are they not all used to find the relationship? How do you choose a specific RCM simulation? You should also remind the reader at the start of this section that this is done at a grid box level. Figure 3: Could you zoom in to show what the red line looks like? It's quite hard to see. Equation 2 must be more complicated than it looks to produce this line. Page 8, Line 5: 'The unitless α' – α is five values obtained from 365 (maybe more for the five RCMs) so I can't picture what it really is or how it represents sensitivity to temperature. I'm sorry I'm quite lost with respect to alpha. Maybe more equations would help. Page 9, Line 6: alpha depends on the RCM. How have you used the RCMs here? Have you chosen one – if so, why start with five? What happens to the data from the other 4? Have you found alpha with all five and then randomly selected from all of them to produce the MC sample? Perhaps reordering the writing here might clear things up with regards to how you are using the RCMs.

3.3 Indirect response. . ..

Page 10, Line 26: What does series mean here? Please be more precise. Page 11, Line 4: I don't understand this point. Can you reword it to be more precise.

3.3.1 Identifying the modes. . .. . .

Page 11, Line 6: Are you referring to the residuals in Figure 3? Your language appears to have changed here and it sounds like you are doing something new. If these are the residuals please use the same language and link it better to the previous sections. Page 11, Line 7: 'Where the patterns of variability obtained from EOF projections of VCsN. . ..' Figure 5: Can you interpret these EOFs? Why is the period 1972-2013 used here? Also, this is not the RCM discussed in the introduction? Why have you changed the RCM?

3.3.2 Modelling forced changes. . ..

Page 13, Line 12: Is it just New Zealand or is it likely to be more representative everywhere? Page 13, Line 14: Equations would be help clarity here. I'm still struggling to understand how you are correlating data on different scales. Page 13, Line 23: How did you remove the autocorrelation? Page 13, Line 25: Why did you not remove autocorrelation at larger time lags? Can you talk about the implications? If it would matter that the interannual variability is too small why not calculate and remove it? If it's not 'too small' then justify not doing it. Page 13, Line 28: How did you create pdfs from time series? Is the data from the whole time series in the pdf or are these time slices? What is the implication of having time series data in these pdfs? Have you got a sample of PCs to create the pdf?

3.3.3 Modelling higher order modes. . ..

Page 15, Line 2: Can you better explain where ïĄş comes from? Perhaps use an equation. Page 15, Line 9: Why is n=50? Figure 7: Can you show a close up of one of the blue lines perhaps – it's difficult to see how the smaller scale temporal patterns are captured and discussed later. Page 17, Line 5: I don't understand the line starting 'The apparent annual cycle. . .'. Could you elaborate? Is this what you'd expect and is it adequate? Page 17, Line 9: Are there any implications to the interannual variability being smaller? Is improving this a future direction? Page 17, Line 23: Can you say where 19000 comes from? Previously, you mentioned 1900 simulations so I'm unsure

what the extra simulations are taking into account.

Technical points: Abstract: Page 1, Line 11: change the direction of the first quote mark. This happens in other parts of the paper too. Introduction: Page 2, Line 7: Do you mean ‘uncertain’? Page 4, Line 28: Make all of global subscript – use _{global}.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2017-36, 2017.