Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-170-RC1, 2016

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Exploring global surface temperature pattern scaling methodologies and assumptions from a CMIP5 model ensemble" *by* Cary Lynch et al.

Anonymous Referee #1

Received and published: 28 August 2016

This study uses two alternative methods of pattern scaling and compares their results.

I regret to have to suggest rejection of this paper. I find it extremely confusing and interspersed with unjustified assertions. I also belief that the analysis is flawed, unless the opacity of the description of the methods used is hiding something from the reader and the authors have actually performed things differently from what it appears. It's been very difficult to follow and guess what is going on here.

I start from a fundamental disconnect in the interpretation of the regression method. The authors claim that after deriving the beta coefficients they go on to compute the 21st century trend by multiplying it by the number of time steps (page 5, line 14). I don't think this is correct. The beta coefficient, by construction, fits the local trend to the

Printer-friendly version

Discussion paper

global average temperature time series, so the 21st century trend is only recovered by multiplying the beta coefficients by the global temperature time series, and then fitting a linear trend to the result. So I start by wondering if the entire comparison here is flawed from the beginning by this misinterpretation of what beta is. After all, beta should allow the user of the pattern scaling approach to simulate any new scenario, given its global average temperature time series derived through a simple climate model, while if one followed the authors' recipe any scenario projection would look the same as long as it takes place over the same number of years.

The statistical analysis of differences using a t-test is also open to discussion. The authors use some jargon about a "2-tailed Student's t-test probability using the incomplete beta function ratio" (page 5, line 16/17). I have no idea what that means. The following sentence is also rather opaque (like the majority in this paper, I'm afraid). I found myself sprinkling question marks to the side of almost every paragraph.

I am doubtful that the comparison of regression and delta method (even leaving aside the problem of misinterpretation of the beta coefficients) is justified the way it is carried out through the paper. Why is the comparison always between the regression method that uses the 21st century and the delta method that uses the 19th century baseline? For cleanness, wouldn't it make sense to use the delta method that uses the 20th Century baseline, which by construction addresses the change over the 21st century rather than the change over the 21st AND the 20th century?

What is the point of including the discussion of reanalysis trends? It confuses the reader in thinking that pattern scaling has anything to do with observations (or reanalysis, for those that do not like to think of the latter as observations).

I will now point out many sentences that I find difficult to understand, or hardly justified.

Starting from the abstract: "Temperature patterns generated by the linear regression method show a better fit to global mean temperature change than the delta method" How is the fit to global temperature change of the delta method defined (and therefore

comparable)? "global mean temperature sensitivity is higher" (in lower forcing scenarios). What does that mean?

Introduction Why start with the RCP discussion and then talk about computational costs of running scenarios (when all RCPs were in fact run in CMIP5?)

"scaled scenarios are used for reducing uncertainty". How is that? In fact, the use of scaled scenarios, one could argue, increases our uncertainty by allowing us to use a larger range of futures than would be available otherwise.

The description of what pattern scaling is and how it relates to SCMs is not clear at all, it never mentions the idea of estimating patterns from available GCM experiments and running a SCM to derive a global average temperature time series, which is the whole point of pattern scaling. A reader less than familiar with pattern scaling would be completely in the dark as to what pattern scaling is and how it works in providing alternative scenarios for analysis.

Assumptions I have big troubles recognizing these assumptions as driving pattern scaling. Pattern scaling by definition does not address internal variability at all. Also, "in practice, estimation errors introduced through this assumption are small". Who says so? This needs a reference.

"In the linear regression pattern scaling method, the underlying assumption is that local change scales proportionally with global mean temperature (GMT) change, and that the relationship is stationary over time". This is actually the assumption of pattern scaling, period, not of the linear regression method uniquely.

"For temperature related variables the assumption of stationarity is valid,..." except just a few lines later you say "when pattern scaling patterns of temperature extremes, the magnitude of the error in the pattern estimates was substantially large". Aren't temperature extremes temperature-related variables? And in any event, the statement "the assumption of stationarity is valid" is open for debate. it may be an ok assumption

according to some error metric, maybe.

"we use a simplified approach for each method to assess the differences in pattern strength and skill" what is pattern "strength" here? And what is simplified (or simplistic, the way it is worded later in the paper, which I find a little negative as a term)?

"potential mitigation" why potential?

Climate Models

"Because models varied in spatial resolution, when appropriate, the models were first regridded to T85 resolution" what does "when appropriate" refer to, here? When was it not appropriate?

Data Analysis

The discussion of the biases in reanalysis affecting pattern scaling methodologies is completely unintelligible to me. I am not aware of any methods of pattern scaling that uses reanalysis or other observational data.

The discussion of different sectors using different periods of reference is again difficult to understand in this context.

The idea that pattern scaling should use an ensemble of models is actually open for discussion: pattern scaling was essentially developed for single models, and work in the literature has shown how high the variability is of patterns across models, so I am not sure that taking this view would satisfy any practitioner of pattern scaling. After all, the most popular pattern scaling tool, MAGICC-SCENGEN provides a library of patterns, one for each model.

"we assume the ensemble variance for each pattern is the same" what does this mean? Couldn't you compute the actual ensemble variance for each pattern and use that in your statistical significance testing?

Epoch differences

Here and in other places a big emphasis is made on the assumption that the trend has to be the same, and I fail to understand why that is the case. The delta method has been applied across scenarios, it has been shown that different scenarios, as long as they are mostly driven by increases in GHGs, produce very similar patterns, but the trends in different scenarios are far from the same. So it would not seem to be a requirement, and in fact it cannot be, given that the method is developed to produce new scenarios, most of which will differ in trend.

The discussion of variance between the two reference epochs is completely obscure. I don't actually understand what variance we are talking about. That of models? That of temperature itself? And how is that relevant here?

Pattern differences

"When using the lower forcing scenario, it is not necessarily implied that there is less future variability" this is a sentence whose meaning in this context is not easy to understand.

Is that 0.8 on line 25 of page 7 supposed to be 0.08?

Scenario differences

The discussion of how lower scenarios produce less change in GMT and that explains less variance explained may be justifiable if you discuss the role of internal variability, but you don't.

"This is due to the fact that scenarios with stronger mitigation practices have a smaller GMT trend and the resulting local temperature sensitivity to GMT is stronger" what does this mean, why should it be the case?

Conclusions

"Choice of scenario can affect the resulting pattern, particularly when using the regression method"

I can't help wondering if this claim is the result of what I think is a mishandling of the beta coefficients.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-170, 2016.