

Earth Syst. Dynam. Discuss., referee comment RC3
<https://doi.org/10.5194/esd-2022-14-RC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on esd-2022-14

Anonymous Referee #3

Referee comment on "STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations" by Claudia Tebaldi et al., Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2022-14-RC3>, 2022

Review of "STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations" by Tebaldi et al.

This paper presents a procedure to create surrogate trajectories of climate model ensembles. The authors provide tests on a set of CMIP6 simulations and discuss the sensitivity to two key parameters of the procedure.

I have no reason to doubt that the authors know what they do. My main concern with the paper is that I neither understand the general picture nor the details.

My first concern is on the format of the paper and its suitability for ESD. The abstract, introduction, and conclusions are written by and for IPCC insiders, as the authors use a lot of IPCC jargon, which is obscure to most human beings, including me. This style of writing seems to go against the interdisciplinary nature of ESD. Not only the paper does not report new understanding of the climate system, but the authors do not discuss that their procedure might help do so (or how). Another example is the use of the term "emulator" or "emulation". Of course, this remark is not limited to this manuscript. I yet have to see a reasonably clear definition of what is called a "climate emulator". For some authors, an emulator is a regression between some predictand variable and a predictor. Here, it is obviously something else, that looks akin to analog modelling. Making a proper bibliographic search could help relate the procedure described in the manuscript to existing work, which might not appear in the IPCC reports. The notion of "creating new scenarios" is not clear. The IPCC seems to use SSP scenarios, which are relevant for the economy. What the authors do is obviously something else. So, using this terminology might be confusing. The simple (acknowledged) fact that the emulation procedure cannot produce relevant GHG (or any forcing fluxes) should plead against the use of "creating scenarios". My understanding is that the procedure creates surrogate trajectories that are constrained by GSAT values. Why should those trajectories be called "scenarios" in the IPCC sense?

My second concern is that the procedure description seems inappropriately vague. Ideally, I should be able to reproduce the procedure by reading the manuscript (provided I have access to the data). The first step (l. 148) suggests that *one* time series of GSAT is created for each model by dumping together all ensembles, scenarios, etc. ([...] "the time series is made [...]"). I guess/hope that the authors do differently. The fourth step (l. 157) is not clear: what is a target scenario? The authors allude to "target scenarios" in several places, but do not define what those are. I believe that the authors could design a diagram that explains how the procedure works. In practice, I understand that one needs to know the target scenario (i.e., have GSAT data). Hence, I do not understand how the authors can reconstruct "unknown" scenarios (e.g., SSP2) from just SSP1 and SSP5, which suggests that intermediate scenarios can be deduced from two extreme SSP scenarios. This might be true, but I would like to understand this miracle (at least for me).

My third major concern is on the results or the performance tests. The authors seem to be happy with the results reported in Figs 1-8. Indeed, the "emulated" time series are close to "targets" (whatever how the targets were designed). But is this desirable? The GSAT time series have no decadal or interdecadal variability (which might due to the procedure itself). This is not discussed, but I would doubt any procedure that creates trajectories that do not yield long term variability are so useful, or really account for climate variability (e.g., the so-called butterfly effect). For me, the SOI results are "good" by construction, since they are excerpts of existing simulations. How would this emulation procedure be able to emulate changes in ENSO variability, which would be a key issue for impact modelling? My feeling is that the simulated trajectories give overconfidence about (the lack of) climate internal variability. The conclusion that this procedure can replace numerical model simulations hence seems overconfident.

Minor issues

In the search of nearest neighbors in the (T, dT) space (step 5, l. 165), are there different weights on T or dT, in the distance definition?

In step 6 (l. 170), what is a "pointer"?

l. 211: "(see 1)", what is "1"?

Figure 1: I can't read the labels on a printed version of the manuscript.

I feel that there should be a separate section that describes the experimental set ups, tests, etc.

Figures 2-4: the captions should only keep descriptive statements, not comments that already appear in the text.

Equation (1) (l. 360): all symbols should be introduced. What is the bar? I think that \hat{y} should be the synthetic and y the truth, not the other way around, as suggested in l. 362. E_2 is certainly not a ratio of variances (but a ratio of standard deviations). The denominator of E_2 should be: $\langle (y - \bar{y})^2 \rangle$ (no $\hat{}$).

In conclusion, my feeling is that the manuscript would be much more appropriate in GMD, which incidentally has a better impact factor than ESD. Of course, this decision is left to the authors and the editor.