PONE-D-18-14400 Propagation of Error and the Reliability of Global Air Temperature Projections PLOS ONE

Dear Dr. Frank,

Thank you for submitting your manuscript to PLOS ONE. After careful consideration, we have decided that your manuscript does not meet our criteria for publication and must therefore be rejected.

Specifically:

First of all I would like to apologize for the long time that it has taken to reach a decision on your manuscript, but as I am sure that you understand this has been a consequence of trying to get a fair outcome.

The work that you submitted for consideration is challenging for several reasons. First you acknowledge that your work has been rejected for publication in other journals with up to 27 reviewers having checked previous versions. As an Editor this is something that I appreciate, being honest and transparent about the study. Secondly you request that we ban from the review process some of the more respected experts in the field of your work and what is more, you suggest that only experts in one field have got the understanding to judge your work. This complicates the assessment of your work. Finally your manuscript questions very well established and broadly accepted research results. Going against the consensus is not easy and therefore your previous requests are partially acceptable.

I have to say that despite all of this your letter sounds slightly arrogant. Being very open about this, I do not share your view and I can not accept your statement saying that for example, I do not have the capacity to assess your work. I have got a BSc, an MSc and a PhD in Physics, with specific courses on the study of error measurement. On one of the shelfs in my office I have a book dealing specifically with accuracy and reliability and I have cited it in some of my works. The fact that people like me has specialized on Earth Physics does not mean that we do not have extensive training on radiative balance, chemistry, etc. You should have it into account, because when submitting a work, to begin a letter saying that your are the smartest person in the room could condition an editor or reviewer. Now I would like to make clear that this has not impacted my evaluation of your manuscript.

That said, there are several reasons why I consider that your work is far of being publishable in its current form. Some of them are more 'formal' and others more 'scientific'.

Beginning with the formal points:

- lack of focus on the topic that you address: from the title and abstract your manuscript is about error propagation in climate models, however in a 50 pages main manuscript, the first 25 have nothing to do with it. This 25 initial pages should be a different manuscript itself. They are about a completely different topic, computation of the radiative balance in the troposphere. There is no point on presenting all this material together

- bad organization: your manuscript has a very bad balance, only 3 sections in the main part and then you go up to 10 subsections that you actually need to make the point sometimes

- lines 112-137 are completely irrelevant here

- as an example, in line 273 you cite 56 and 57 and in my view these works do not support your claim, indeed they could be interpreted as serious criticisms to what you expose.

Scientific points:

- great claims need great support/evidence: basically you state that the IPCC is wrong. The point is that if you want to prove it (and to go against the broadly accepted consensus) you need to refute at least some of their claims, and in your work you do not do it. I can accept that it could not be fair from the formal point of view to request you to demonstrate that others are wrong, but at least you should strut your results with demonstrations of the style 'and this probes that the result obtained in previous works [X] is wrong because of this and this and this';

- you do not address a big point in all the debate. For example, empirical observations all over the world evidence that despite any issue with error propagation the method of using ensemble mean works. If we assume that your work is ok, at lest as a summary in your conclusions for a well balanced discussion it is necessary to state something like 'despite all these problems and limitations over the last 30 years the models are proved to be useful and to be able of forecasting with a very high skill in 1980s the mean climate for today' (empirical evidence);

- I do not agree that we are facing 'physical' errors as you claim. Any deviation from reality is not because of the physical system here, it is because of the tool and measurements and completely independent of the level of understanding of the completeness of the theory. They are mathematical/statistical deviations. In a text dealing so much with accuracy making clear this difference is important. Other fields of study are also dominated by formalism without analytical solutions and not because of this those using results projected consider they wrong. Also the error of a measurement can be known partially a priori because of the instrumental error and this does not mean that the result is considered 'erroneous' in any discipline or field. Here you should be aware that in some cases your language use is excessive and this does not help to consider the presentation of your results fair: 'erroneous', 'mortal test', 'negligible CO2'...

- your introduction to parametrization in lines 141-143 is 'obscure'. It is not clear to me what you want to say. What do you mean by not empirically tuned. Of course some parametrizations are empirically tuned, this is why they are parametrizations and not part of the free running of models;

- another issue comes out with your assumption about the cumulative nature of the error. This could be a discussion/paper itself. Part of this is to accept that the mean climate (or let's say the total cloud fraction) for the step 't+1' depends only on 't' and probably this is not true and it is completely model dependent. There are a lot of other factors that can contribute to TCF values in a model, beginning with a fully coupled land model with biogenic particles, statistical values of

parametrized volcanic eruptions, number of chemical species interacting... Why should we assume such additive nature in a model where boundary conditions at t are different than those in t+1?

- and just to add one more thing, for example, in lines 616-617 you use a reasoning pretty weak: if you argue that the fact that something works does not mean that it is ok, then you can not admit that statistical improbability is scientific proof of the opposite argument being true.

To summarize, obviously my assessment includes much more material that I have written here and these are just some picks representative of what your work includes and I have had into account. However because of all these concerns with the lack of focus on the presentation of your manuscript, unbalanced discussion and lack of evidence to support part of the results that you present, your manuscript can not be accepted for publication. For future works I would suggest you splitting this work at least in three different parts that I think are pretty clear.

I am sorry that we cannot be more positive on this occasion, but hope that you appreciate the reasons for this decision.

Yours sincerely,

Juan A. Añel, Ph.D. Academic Editor PLOS ONE

[Note: HTML markup is below. Please do not edit.]

[NOTE: If reviewer comments were submitted as an attachment file, they will be attached to this email and accessible via the submission site. Please log into your account, locate the manuscript record, and check for the action link "View Attachments". If this link does not appear, there are no attachment files to be viewed.]

For journal use only: PONEDEC3