

Interactive comment on “Agnatology: learning from mistakes” by R. E. Benestad et al.

M. Huber (Editor)

huberm@purdue.edu

Received and published: 3 September 2013

Since this submission has led to a wide and vigorous debate the editor has decided to include his reasons for not recommending publishing the paper in ESD. The decision follows.

Based on the reviews and my own reading of the original and revised paper, I am rejecting the paper in its current form. The submission is laudable in its stated goals and in making the R source code available, but little else about the paper works as a scientific contribution to ESD. While I think as an ESDD publication at least a discussion was had and the existence of the R routines has been brought to the attention of the various interested communities, the manuscript itself is not a good fit for this journal and would need substantial further revisions before being ready (if ever) for this journal.

The problems are several fold.

C400

First, I do not think the structure of the paper works. The long, didactic introduction is not appropriate for this journal and all the meat of the paper is currently in the appendix which is a strange place for it. Indeed, as currently structured there is no paper in this paper, i.e. there is no actual science (hypothesis, testing of a hypothesis) in the main body. The historical lessons and systemization of error may be scholarship, but not in this (ESD) field and may be more appropriate for a different audience (I'm thinking Physics Today or a philosophy of science journal).

Second, much of the discussion in the appendix is written in an inflammatory and insufficiently supported fashion. Removal of subjective characterization would make the paper stronger by reducing the verbosity and of more lasting value by focusing on scientific issues. It is entirely irrelevant whether the authors of some papers also distribute pamphlets to school headmasters, just as it is scientifically irrelevant what the political affiliation or religion or hair color of authors are.

Third, while much is made that so-and-so made mistakes, much of that characterization relies solely on the authors' stated opinion. While I agree that demonstrating how results may differ based on various choices with the R routines is useful, it generally (except in the case of coding errors) does not reveal mistakes. Instead it reveals how different choices lead to different results. It is really up to individuals and communities to determine that something is a mistake (or something that otherwise contributes to continued ignorance). Let me emphasize this point since it goes to the heart of this paper. I see very little in this paper that actually demonstrates real flaws in prior work. Instead, mostly we are dealing with flaws of type B and C (in the paper's nomenclature). In fact, I would argue that a number of the issues classified as flaws of A and D type are really just flaws of B type in disguise (what statistical tests and signal processing tools are used is largely a matter of the norms and history of the field in question, hence multi-disciplinary work will always lead to the appearance of 'incorrect' analysis by members of one or more communities). Flaws that arise from an incorrect logical premise sound straightforward to identify but may be harder in practice to nail down

C401

then the portrayed in this manuscript. I'm not being a relativist here, but I think the paper dances around the main issue being raised by the various authors (and that appear in their commentary online).

The root logical flaw in many of the papers discussed in the appendix is that showing a statistical correlation between some non-CO₂ variable and some observed climate time series somehow disproves the hypothesis that CO₂ is a driver of climate change. This is as silly as saying the cost of my sneakers is correlated with how fast I run and therefor I have invalidated the hypothesis that training makes me run the 100 yard dash faster. Do we really need 70 pages of text and two dozen R routines to recognize the logical problem here?

And therein lies the real problem. The climate science community has strong theory (dating back more than a century) and good, physics-based models that underly the attribution and prediction endeavors and these guide the interpretation of observations and their statistical characterization (i.e. what the null hypothesis is). If one ignores that foundation as most of the studies being criticized in this submission do, then one is left with unconstrained statistical analyses or curve fitting exercises that have no clear plausible, physically viable explanation. The reality is that many of the authors whose work is being criticized are on the record as thinking that either climate theory and/or climate models are fundamentally flawed, hence the adopt the kind of approach which leads them to conclusions that are in opposition to the vast majority of climate scientists. Again, this can be said in two sentences.

A stronger paper would show exactly how what we know based on first principles physics and more sophisticated models guides the choice of null models, statistical tests, and curve fitting techniques, and how such physically-informed choices would alter the conclusions of various previously published papers. Rather than concentrating on ignorant mistakes made in prior work, it would help to identify the key parameters in those studies and justify *a priori* what the value of those should be. A fair amount of that is in the submission, but it is buried within invective and verbosity (or within source

C402

code).

I would entertain a new submission with a complete reworking along the lines I have indicated and I would send it out to new reviewers and there would be a fresh comment and response period. Or submit elsewhere and hopefully the paper has been improved through the ESDD processes.

-Matthew Huber

Interactive comment on *Earth Syst. Dynam. Discuss.*, 4, 451, 2013.

C403