Review of
AGNOTOLOGY: LEARNING FROM MISTAKES
by Benestad et al.

1 Scientific Comments

Replication is fundamental to the scientific method, and I agree with the authors that it is an under-rated quality in the scientific literature. However, it does not explain all that is wrong about the fallacious studies they investigate. While the multiple bogus cases dissected by the authors all need to be formally rebuked in the peer-reviewed literature, I am not convinced that it should be done as in the present manuscript, under the somewhat pompous banner of "agnotology", or under the premise described in the abstract. I expand on these points below:

Philosophizing The introduction is needlessly philosophical, obscuring fairly simple concepts un-
der a verbose cloak of lofty assertions (with many typos and grammatical mistakes to boot!).
In my experience, there are three main reasons why scientific studies might be flawed:

1. One may start from a correct logical premise and execute an erroneous analysis
2. One may apply an correct analysis but start from the wrong logical premise
3. One may start from the right premise, and correctly apply the analysis, but overstate the significance of the conclusions (i.e. the analysis does not actually address the question)

The paper mainly deals with #1, leaving the other two possibilities somewhat buried within the laundry list of the appendix. The two cases selected for prime time do not seem to constitute enough material for a paper. It seems to me that the authors face a choice: do they want to write a philosophical opus? If so, they should make sure to cover all logical possibilities for scientific ignorance or lack of consensus, and send it to the Transactions of the Royal Philosophical Society or a similarly-minded journal. If they want to write a paper on reproducibility, they should make this focus plainer. As it stands, the paper is an unwieldy hybrid of philosophy of knowledge, replication studies, and scientific meta-critique.

Overfitting I think the main problem with the two cases presented here is the logical fallacy of
fitting a statistical model to a set of observations, using a the calibration skill as an implausibly
rosy measure of prediction skill, and going on to make predictions about the real system
without making sure that the model has adequate out-of-sample skill. This is simply known
as overfitting, an elementary statistical notion that is explained in any good statistics textbook.
It is sadly common in science, usually because non-statisticians are unaware of the problem,
and happens usually because of ignorance rather than an intention to deceive. Perhaps the

1

authors could use this framework to kill several birds with one stone, as it is a mistake common to many of the papers they scrutinize? It would seem aptly philosophical to subsume many studies under one common framework rather than segmenting the issue into many small pieces.

Statistics vs logic The basic premise of the paper is that many of the errors in the denialist literature indicted here stem from methodological flaws. However, in the case of the multiple bogus studies by the prolific Dr Scafetta (whose work has been debunked in many places, including *Benestad and Schmidt* [2009] and `http://agwobserver.wordpress.com/anti-agw-papers-debunked/`), the issue is less the analysis itself (though there are certainly plenty of problems there), than the underlying logic. However, nowhere in the main text is the word "logic" ever used, which I find odd.

Scafetta [2012] purports to test climate models over the instrumental record, ignoring the fact that no GCM is ever expected to match observed temperature, in light of natural climate variability. In doing so the author sets up a strawman, an impossible task that models are not designed to achieve. In contrast, overfitting a sets of sinusoids will, by construction, match the training data, but may radically mis-predict out of sample observations. The logical fallacy here is that a model with no physics will outperform a physically-based model, not understanding which aspects of the climate system the performance should be judged on. In this case, it is fair to ask the models to reproduce various statistics of temperature observations (e.g. mean, power spectrum, higher order moments), but not the *phase* of these variations, which may be viewed as random.

In many of the cases they investigate, the authors allude to these logical flaws as "unclear physical basis", which is a little too vague. I think "agnotology" would be better served by precisely identifying logical fallacies.

2 Minor Comments

- The English is an unholy mess, which is shameful given that there are several native English speakers amongst the authors. This is just not serious and makes the reader feel like their time was wasted. I normally point out all the typos and grammatical mistakes, but in this case there are too many to type in a review, and given the comments above I don't think it would be a productive use of my time. I do have a manuscript covered in red ink, if the authors are interested!

- The wikipedia page on agnotology claims that the first occurrence of the term dates back to Proctor, Robert N. (1995). Cancer Wars: How politics shapes what we know and don't know about cancer. New York: Basic Books. ISBN 978-0-465-00859-9. I have no access to this book but I recommend that the authors check this, as it would change the date of the terms first use (L20) by 13 years.

There is every sign that climate ignorance amongst the public is (like the tobacco scandal) due to a malicious intent to deceive, perpetrated by scientists who knowingly produce junk science in the hopes of confusing public [*Oreskes and Conway*, 2010]. Exposing those logical and methodological flaws is a valuable endeavor aiming at dissipating this cloud of doubt. While the cases investigated here are quite interesting, and deserve to be published in an open scientific journal like ESD, I cannot recommend publication under current form. Indeed, the title, introduction and discussion of the paper appear to be a hastily-written wrapper slapped together to package these examples

for publication. The fact that those sections are so ill-written only serves to weaken the arguments made in the 17 cases presented here. I recommend that the authors restructure their paper around common themes (e.g. logical fallacies or common methodological mistakes), write it clearly and concisely, avoid snarky comments against denialists (irritating though these characters might be!) and work on a coherent presentation, instead of publishing a laundry list of replication studies and wrap it in dubiously written philosophical verbiage. This will make for a much stronger contribution to the scientific literature.

References

Benestad, R. E., and G. A. Schmidt (2009), Solar trends and global warming, *Journal of Geophysical Research (Atmospheres)*, *114*, D14101, doi:10.1029/2008JD011639.

Oreskes, N., and E. Conway (2010), *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*, Bloomsbury Publishing USA.

Scafetta, N. (2012), Testing an astronomically based decadal-scale empirical harmonic climate model versus the ipcc (2007) general circulation climate models, *Journal of Atmospheric and Solar-Terrestrial Physics*, *80*, 124–137.