Review: Bayesian deconstruction of climate sensitivity estimates

This very well written paper attempts to argue for a more transparent approach to quantifying uncertainty in climate sensitivity estimation by using explicit statistical models and justified priors within Bayesian analyses. I agree entirely with this position and so there are things that I like about this paper. However, I believe that the core underpinning assumption that the authors make, namely that interval estimates given by researchers on observations should be treated as likelihoods within a Bayesian framework, cannot be justifiably imposed on a researcher and makes little sense even if it could be. I will detail my objections below and suggest a correction that would enable calculations to be redone and the paper to be resubmitted. I am therefore recommending that the paper be rejected and, if the authors wish to make alterations to their methods and argument, resubmitted for publication. I want to stress that I like the concept of the paper and the idea to examine climate sensitivity estimates as if certain priors were chosen is a nice one, so I genuinely would hope that the authors do resubmit having reflected on the comments below.

Major corrections

The critical problem in the paper's argument is as follows: Assuming, using the notation of the paper, that x_T is the true temperature and x_o is an observed temperature, then the measurement model is, as given,

$$x_o \mid x_T \sim \mathrm{N}(0, \sigma^2).$$

The likelihood of any particular measurement, x_o is then $p(x_o \mid x_T)$ and has the form given on page 3 line 20. The likelihood is a function of the unknown and unobserved x_T. On line 10 of page 7, the authors make the assumption that is key to the paper's methodological argument, that published estimates for ΔT, should be correctly understood as representing likelihoods $p(\Delta T_o \mid \Delta T_T)$. This cannot be true as is, because the likelihood is a function of the unknown ΔT_T, and an estimate is a given interval. However, the authors clarify their meaning in the second half of the sentence as meaning that this is a 'likelihood' with "an a priori unbiased error of the specified value", which I take to mean that they assume x_T was given an unbiased estimator (not Bayesian) and then plugged into the likelihood. But the algebra doesnt match the description. Line 11 says that $p(\Delta T_o \mid \Delta T_T)$ "provides an uncertain estimate of the true value…", but the true value of Δ_T is considered known in the given conditional.

The described procedure does not define a likelihood. You might call it the likelihood function at $\Delta T_T = \Delta T_o$, but the authors want to do that in order to claim that researchers have simply adopted a likelihood and then committed the fallacy of the inverse. It is not clear why you should impose that constraint on researchers who have given an interval estimate, particularly when a more natural Bayesian interpretation of an interval is available (see below). Before explaining, it's worth examining the author's claim in the context of their example on page 11. Line 11 states that "we recognise their estimate $\Delta T_o \sim \mathrm{N}(0.77, 0.08)$ as a likelihood $p(\Delta T_o \mid \Delta T)\ldots$". So, the idea is to take a researcher's distribution for ΔT_o, given unconditionally, and assume that it's derivation has committed a fallacy because they really supposed that they knew ΔT, the key unknown they'd like to estimate, and fixed it at ΔT_o. The authors are correct to point out that this would be problematic, but it is not a natural interpretation.

When we see a distribution such as $\Delta T_o \sim \mathrm{N}(0.77, 0.08)$ or an interval estimate $\Delta T_o \pm \sigma$, the authors are right that we cannot simply suppose that the same distribution holds for ΔT, but that is not to say that we can impose implicit conditioning on ΔT either. We have a distribution for ΔT_o and nothing else, and I don't see why we should be able to argue to rewrite the claims of other scientisits to fit a given narrative. A more natural Bayesian interpretation and one that does not assume that "scientists who wrote X really meant Y", is to take the distributions given at face value, so suppose we have been given the distribution for ΔT_o that the researchers believe. This is the evidence or the marginal likelihood in the Bayesian paradigm:

$$p(\Delta T_o) = \int p(\Delta T_o \mid \Delta T) p(\Delta T) d\Delta T,$$

where the likelihood given by the usual measurement model is integrated over the prior for ΔT.

As this preserves the given intervals and still enables the authors to take their approach in investigating implicit priors and changing them to more natural ones, a revised version of the paper should do this. It may be that the description of the implicit conditioning resonates with certain scientists who publish interval estimates, so the authors could keep some of their originial analysis, but more carefully caveated.

But even the above correction, still assumes that the original researchers were subjective Bayesians (or at least ascribed to a subjective interpretation of probability). A difficulty I have with interpretations of analyses from one philosophical school (e.g. frequentism) in terms of the others (e.g. subjective Bayes), is that they are fundamentally incompatible and a great deal of straw manning the other side is required to obtain a coherent argument. I am a subjective Bayesian, and I would argue that Bayesian approaches are the only ones that make sense in any given situation. However, if a researcher has undertaken a frequentist analysis, I would be very careful not to overinterpret. A key example here is alluded to on line 9 of page 5. If uncertainties in measurements are purely random errors coming from a measuring device, the σ in the measurement model has a clear frequentist meaning (up to the debate about what purely random means, wherein Bayesians and Frequentists are likely to also disagree). A frequentist analysis can admit no further type of uncertainty here, nor would they mean to when they reported a confidence interval for x_T. It doesnt even make sense to talk about a probability distribution for x_T, as x_T is not random, and a good frequentist would not make the mistake that it was. A Bayesian might also consider contributions to uncertainty through assumptions required to compute the derived observed quantity and other things that are unknown, but not random. The very interpretation of the measurement model and what is/can be included in σ can be different, and it makes sense to talk about a probability for x_T, albeit a subjective one.

The fundamental problem in a lot of the mentioned publications is that the statistical model is not clearly written down, the assumptions are not clearly stated and the meaning of uncertainty statements is not clearly given. This leaves a void in which readers are free to project their own definitions of "uncertainty" and where the meaning of inferred distributions or intervals is open to interpretation. The authors here have recognised this and attempt to give some interpretation, but there is a danger that with any of the analyses they picked (perhaps with the exception of their own), that they are straw manning the position of the authors. A safer way to do what the authors want to do is to make statements such as "if paper X meant Y by this statement, then that would mean Z, but A, B and C are also possible interpretations". It's harder work of course, but, that is the problem with traversing the different philosophical positions, or, more precisely, with attempting to ascribe subjective Bayesian interpretations to existing analyses that are unlikely to have been conducted by researchers who hold those views. Aside from frequentism, there are also at least 3 kinds of Bayesian (subjective, objective, as the authors mention, and falseificationist), and each would mean a different thing by the modelling statements and conclusions that were made. I like the Bayesian approach that is taken in the paper, and I see no reason not to advocate for it. I do believe that the set up and the discussion and interpretation of other work should bear the above in mind.

My final point is on the first 4 pages where Bayesian statistics and the prosecutors fallacy are exaplined in far too much detail. Bayes is widely used in climate science, so I think a derivation of Bayes theorem (seen in any first year undergraduate probability/statistics course) is overly indulgent. Similarly, the prosecutor's fallacy is well known and is taught in every entry level probability and statistics course, so a page and a half of a research article should not be given over to explaining it in detail with examples. A quick outline and references should be enough.