Answer to Short Comment #1 by Nicholas Lewis

We thank the author of the comment for their constructive criticism and we respond to the points in the following, where C: is a comment brought by Nicholas Lewis and A:, our answer.

C: The paper's criticism of the Kalman filter method (section 2.3), as implying – very likely unrealistically – that the model ensemble is a credible predictor before consideration of the observational constraint, almost ruling out posterior estimates outside the model range, is valid and in my view sufficiently important to warrant mentioning in the Abstract.

A: This, indeed, is a relevant point. As pointed out by Reviewer #2, the Kalman Filter could appear as a significant rival approach. However, we consider it too restrictive, for the reasons reminded here by the author of the comment. Nevertheless, the main scope of the study is not to criticise the Kalman filter method, but to present a Bayesian method which we think is more appropriate to the question of emergent constraints. Therefore, we do not consider adding references to the Kalman Filter in the abstract, mainly to avoid an overload of information that could mislead the reader.

C: However, a major weakness of the paper is that it fails to investigate, or even acknowledge the existence of, an objective Bayesian method that has been applied for a very similar purpose, or of the frequentist likelihood ratio method that has also been so applied (Lewis and Grunwald 2018). Objective Bayesian methods use a 'noninformative prior' that reflects how the expected informativeness of the data about the parameter(s), derived from the likelihood function, varies over the parameter space, and where not all parameters are of interest may also reflect the targeted parameter(s). There is a huge statistical literature on objective Bayesian methods, as there also is on likelihood ratio methods.
Both the aforementioned objective Bayesian and likelihood ratio methods generate un- certainty distributions and ranges that that have been shown, in a perfect model test, to be well calibrated for combining, as well as evaluating separately, independent evidence (Lewis 2018). That is, the uncertainty ranges output by these two methods, although different in statistical nature, are both close to exact confidence intervals. Accordingly, in the long run probabilistic conclusions by an investigator employing either of these methods will on average be true statements, which is surely highly desirable for scientific investigations. That is not in general the case for subjective Bayesian methods (Fraser 2011, Lewis 2014).
Moreover, Bayesian updating does not in general produce satisfactorily calibrated inference when combining evidence, even if the related Bayesian inference from the separate pieces of evidence is well calibrated (Lewis 2013, Lewis 2018). Nor is Bayesian updating satisfactory as a method of incorporating probabilistic prior information, which can however be incorporated under the aforementioned objective Bayesian method. The appropriate way to do so is by treating the prior information not as a prior density to be used in Bayesian updating, but as equivalent to a notional observation with a certain probability density, from which a posterior density has been calculated using Bayes' theorem with a noninformative prior (Hartigan 1965).
In order to achieve satisfactory inference about climate sensitivity when combining evidence, climate scientists need to move on from fundamentally flawed subjective Bayesian methods, and to cease ignoring the existence of objective Bayesian and frequentist (profile) likelihood ratio based methods that are both demonstrably superior.

A: Besides the title of "objective" Bayesian method, which we find confusing and misleading, there are several reasons why such methods are not investigated nor acknowledged in this study. One of the first reasons would be that this paper does not aim at being a summary of every possible Bayesian method, but solely introduces one method for the question of emergent constraints in comparison to other (non-Bayesian) methods used in the past.

Having said so, "non-informative" prior such as Jeffrey's prior, are, in fact, very informative when dealing with a single problem which carry information by itself, such as the relationship between Sensitivity (S) and Temperature (T) in a defined ensemble of climate model. Actually, we do consider that informative priors are a valuable advantage of Bayesian methods (or updating), as it carries the knowledge, with a certain uncertainty, of the original problem (in that case, the plausible range of S). There is no reason for thinking that one prior would be more non-informative than others in every case - priors are more informative than others based on the problem. In the case of this paper, we could consider that the uniform prior is more informative than the Cauchy prior towards high S. However, such affirmation could be completely different with a different set of climate system parameters. Thus, there is no reason for thinking that one specific Bayesian method would have more or less flaws than another.