

Interactive comment on “Empirical Bayes approach to climate model calibration” by Charles S. Jackson and Gabriel Huerta

Anonymous Referee #1

Received and published: 20 May 2016

Review of "Empirical Bayes approach to climate model calibration" by CS Jackson and G Huerta

General recommendation

The topic of the paper is model calibration and the new thing it seems to introduce is the scaling factor S . While this is of general interest for GMD readers, the paper should not be published in its present form. I have several reasons for this:

1) It is not clear to me what is new in this manuscript relative to the older one by the same authors (Jackson et al, 2008) where the scale factor is already introduced.

C1

2) The reason why this remains unclear is that the paper is so poorly written that almost everything remains unclear. In fact, I stopped reading after section 4 because at that point I still did not have any clue about what the authors intend to develop.

3) The mathematics used lacks clear definitions (see below) which is probably the main reason why much of the paper is non-understandable. I found that I was always guessing what the authors intended to say. Furthermore it seems that there are inconsistencies (see below).

4) References to similar or related work, in particular to introductory texts on model calibration, are missing. There is no mentioning, let alone discussion, of previous work (except from the same group), no comparison with other approaches, no pros or cons.

For all these reasons I suggest to reject this paper.

Major comments

1) Although background material on model calibration is provided in Sect. 2, it does not suffice to my feeling. Readers without experience in model calibration want to have a more basic introduction or alternatively references to more basic introduction.

Certainly, there is some literature on this topic and it is not good that (almost) nothing is cited in the introduction. It should be stated what is novel, different, better, etc., in the presented method in comparison with other approaches.

2) The mathematical introduction in section 2 is unclear in various respects and insufficient.

a) It is unclear whether d means either a quantity like "temperature", or a specified value of that quantity, as "300K". In the latter case, is it the "true" value of the observable or a measured datum subject to measurement error? Accordingly it is unclear whether \hat{x}

C2

is an estimate of the true value or simply a model result for an observable d , that differs from the measured d . In the first case, the residuum ϵ can be viewed as a random quantity, but in the second case it is given as $d - \hat{x}$ which in turn are both given as well, thus ϵ should be non-random in that case. I guess that the probability space is the space of parameter vectors m , but that is not stated as well.

b) If I accept that ϵ is a gaussian random quantity, I still do not see why its expectation value should be zero, in particular when there is missing physics in the model which could easily produce a bias.

c) As the meaning of both d and x remain unclear, it is unclear what eq. 1 actually means. Is it, as a function of parameter vector m , the probability that \hat{x} comes out as d ?

d) To my view the chain of arguments gets broken where the authors mention that the covariance matrix can be rank deficient. Here they enter into a side topic (eofs) that does not lead to the goal, which I think is to demonstrate how covariance leads to problems in model calibration. And to my opinion, the latter goal is not reached at all.

e) Finally, if $\epsilon \sim N(0, \sigma^2)$, then of course $\epsilon^2/\sigma^2 \sim \chi^2$ (equation 6). In this sense, the argument under the exponential function of a gaussian distribution is always χ^2 distributed. In a similar way one could say that the log of ϵ is log-normally distributed. It is true, but I do not understand why this is stated here. Further I do not understand why the mean and variance of the χ^2 are important instead of the mean and variance of the gaussian.

3) Section 3.

a) What is a "climate model evaluation matrix"? As it may be arbitrarily defined, how is it related to the covariance matrix in section 2?

b) 2nd sentence: I don't see the argument. I understand that information on model matching data can be used to determine parameter uncertainties, but how does this

C3

statement follow from the first part of the sentence, namely, that statisticians often use a scaling factor in calibration?

c) Please explain how the likelihood function of eq. 11 can be a gaussian although S is not considered a constant. If it is not a gaussian, then it is questionable whether a gamma distribution for S is still a conjugate prior.

d) How can it be explained that the covariance matrix is suddenly reduced to a diagonal matrix in equation 13. Is this because of the EOF transformation? How can the reader see this in equations 12 and 13?

4) Section 4.

In Section 2 k_e was introduced as the effective degrees of freedom, following from an EOF decomposition of the modelled fields. This gives the reader the impression that the determination of k_e is relatively straightforward. Now, in section 4.1., nothing remains clear. It is not clear where the EOF decomposition is in this derivation. The factor $A/2$ is introduced seemingly without necessity, because it is already gone in eq. 19, just after its introduction in eq. 18. It is not at all clear why and how the number of experiments affects k_e .

Minor problems

1) Why is there a section 2.1 when there is no section 2.2?

2) A square root is missing in eq. 2.

3) Misuse of the "=" sign in eq. 6 and in eq. 18

4) line 99: replace phrase "probabilities are narrower".

5) Check brackets in eq. 11.

C4

