

Notes on Bhatti et al

Uncertainty in aerosol effective radiative forcing from anthropogenic and natural aerosol parameters in ECHAM6.3-HAM2.3

This study provides an examination of Effective Radiative Forcing (ERF) uncertainty in the ECHAM6.3-HAM2.3 (hereafter called EC-H) model using a PPE paradigm to examine its sensitivity to a number of uncertain, but important, parameters that govern processes that influence aerosol-cloud, and aerosol-radiation interactions, and aerosol emissions. A Gaussian process emulator is used to expand the number of members of the PPE in order to more accurately evaluate the parameter values entering into the comparison. The results are compared to POLDER-3 data (AOD, SSA and AE). A summary is provided of the important (and less important) parameters driving uncertainty in this model, and those results are compared to similar studies from other models.

I am a climate modeler with a pretty broad range of experience in model development and use, but I do not have as much expertise as some authors of this study in single, specific areas, so my comments are going to be quite general, rather than focusing on specific items in the study. The study has been done carefully and thoroughly, and I have no specific requests for corrections in methodology, or disagreements with conclusions. But I do think revision would be useful to deliver messages more crisply, cleanly, and completely to make it easier for readers to extract information and put their results into context more quickly. I also think the discussion of structural uncertainty and how it relates to model uncertainty is somewhat superficial and it could be improved. I had a few specific questions about results in figures that should be addressed. I am recommending that it be accepted pending what I think are minor revisions, and offer some comments and suggestions below.

Specific comments:

1. Lines 1-68: A first reading of the abstract and section 1 still left me wondering what the precise target was for the study. Could you state the goals a little more crisply? On line 6 you say your goal is to “address a gap” (my condensation of your words). If I read your sentence carefully, I think you are saying that the gap is in quantifying the sources of uncertainty. So is the goal to identify some of the processes that contribute most to uncertainty in EC-H? And then compare those results to other models? You also eventually connect some of your results on parametric uncertainty with known structural deficiencies in the EC-H model. Maybe you could say something like these sentences explicitly, and provide hints on how (or whether) this information can be used to reduce uncertainty? I also disagree mildly with your characterization of what

structural uncertainty is, and how it contributes to model uncertainty (more below). I found the discussion of Regayre et al (2023) to have helped me understand what can and cannot be gleaned about structural and parametric uncertainty from PPE studies. Perhaps a little more of that kind of discussion could be inserted here also, since many authors are shared between that and this study.

2. Lines 10-11: I had a somewhat different interpretation of your results than expressed in this sentence. I think your sentence could be read to indicate that the leading causes of ERF uncertainty are associated with parameter uncertainties. I am not sure this is the case because some of your results indicate that structural deficiencies play a very large role in ERF uncertainty also, and you have by no means investigated many aspects of structural uncertainty. I personally believe structural uncertainty plays a very large role. I do think it is true that of the parametric uncertainties that you examined there were many common features with other PPE studies. Could you make your arguments more compelling or revise the sentence to agree more with my statement?
3. Lines 15-16: states “PPEs can reduce some structural model biases through parameter adjustments, but others persist.” Maybe this topic deserves more discussion. Such adjustments might lead to a positive outcome, but isn’t this fixing a problem for the wrong reason with the wrong solution? Presumably one doesn’t like to correct a problem by adding an arbitrary correction of a (possibly) satisfactory parameterization to produce a reasonable result with a model containing a significant structural deficiency. It is unsatisfactory if the goal is to improve understanding and representation of the underlying physics, and it is at best an expedient kludge to avoid some other problem.
4. Lines 43-45: Identifying structural uncertainty with deficiencies in “coding” feels misleading. It could be that a process treatment represents one process very accurately, but entirely ignores another. Is this a coding problem? It is an “understanding problem” reflecting a lack of understanding or some other choice by the team responsible for treating those processes. Can you replace “coded” with “represented”, “treated”, “formulated”, etc?
5. Lines 54-55: The phrase “before observational constraints” doesn’t deliver a clear message. Maybe “alongside observational uncertainties that constrain parameter choices”, or some other phrasing.
6. Line 65: It seems to me that the paper has two main goals:
 - a. Characterize the uncertainties for the EC-H model
 - b. Identify some of the shared parametric uncertainties with other models and PPE studies.

It might be worth adding a sentence or two here making these points before discussing what is going to be shown in various sections.

7. Lines 103-106: Are you partitioning the ARI and ACI contributions to ERF using clear-sky and all sky fluxes, or some more sophisticated method (e.g., Ghan (2013), or an APRP method)? Please specify.
8. Line 179: perhaps insert “study” after the word “each” and change “sample” to “samples”?
9. Line 189: typo “perturbationsc”
10. Line 203-205. These lines make explicit that the study is not addressing structural uncertainty. I believe that previous sections of the text were vague on this point, and I also think that the later discussion of the role of the CDNC minimum do deliver messages that are relevant to structural uncertainty.
11. Lines 195-206: I am quite interested in this discussion of common features, conclusions, and differences between the EC-H simulations and analysis done here and studies with other models. I think the authors did a nice job of summarizing some of those differences, and I hope that there is more discussion like this in the summary sections of the study.
12. Lines 216-220. Are you saying that the biases in marine stratocumulus in ECHAM6 are structural, unlike those in HADGEM, UKESM, and the models analyzed by Shindell? I am unclear about the messages in these lines.
13. Lines 220-223. Do you think that the smaller uncertainties you identified in the marine tropics are just due to the crudity of the convective parameterizations and their ACI treatments in all global models, or do you think that those representations are actually robust and we really understand and are representing those processes accurately?
14. Line 233-305: I like the discussion in sections 3.1.1 and 3.1.2 a lot. Would it make sense to change the title of section 3.1.2 to mention “CDNC minimum as an example of Structural Uncertainty” rather than to list the section with the specific topic of minimum CDNC? That way you have a section dealing with parametric uncertainty and another dealing with structural uncertainty (with a big focus on CDNC lower bounds). As you point out in the first sentence or two of 3.1.2 there are other possible contributions. The lower bound on CDNC just provides a convenient and very easy example to demonstrate the existence of structural uncertainty. It might be useful to list some of the possible explanations for the need for such a limiter (CDNC also depends on estimates of subgridscale vertical velocity and internal circulations within the clouds driving aerosol activation, and unresolved or under-resolved exchanges between cloud free and cloudy air masses at cloud top. Other candidates that occur to me that are driven by radiative cooling, evaporation and drop sedimentation. These are processes that occur to me, but you may have a better list).
15. Lines 301-305: I don’t really understand the philosophy behind the proposed PPE study described on the lines of this paragraph. Are you just proposing an alternate kludge (to

the CDNC lower bound), to avoid dealing with lack of understanding of the important processes? I could imagine employing a Machine Learning or AI method to connect CDNC, aerosol properties, and meteorology/cloud properties, but those words and methodologies are not mentioned on these lines, and it is hard for me to envision how one would use a PPE to improve structural deficiencies, as hinted at on these lines. Can you be more explicit about what you are proposing, or drop the paragraph?

16. Lines 306-320: I am assuming the black circles on each line (box/whisker plot) of figure 4 represent the value for a quantity and region for a particular member of the PPE. If I am correct, could you please state that in the caption? And if that supposition is correct, why do all the ensemble member values for AOD, AE and SSA cluster at the extreme ranges of the statistics, rather than frequently falling within the 25 and 75 percentile ranges? Since the circles never seem to fall within the interquartile range it suggests my stated supposition is incorrect, or there is a problem with the figure. So please revise the text, caption and figure a bit to clarify the meaning of the circles and discussion.
17. Lines 312-314: I want to make sure I really understand figures 6,7,8, and the description of the figure is relatively terse in both text and captions. So I will state my guess about exactly what is being displayed, and please revise to clarify. Panel a is showing the POLDER annual mean for 2010. All fields are evaluated at identical points in space and time. Panel c is showing the mean of the emulator estimates for the annual averaged members of the ensemble. Panel d is showing the absolute value of 2* standard deviation of that ensemble of annual averages about the field displayed in panel c. Panel b is showing the difference between panels a and c. Panel e is showing panel b divided by panel d.
18. figure S2 caption. Is the minimum emission diameter for fossil fuels supposed to be 15nm? The caption indicates it is 25 nm.
19. Lines 450-476. I am struggling to extract clear messages, recommendations, and priorities from this section. You list again some of the deficiencies, and the parameters that influence them, but there is no recommendation about how to proceed in order to improve the situation, or a demonstration that a particular strategy could be used to improve the model fidelity. The section closes by indicating that more parameters could be examined, but there isn't really any indication that good use of the present ensemble of simulations could be used to improve the model or the model fidelity. Other than documenting the contribution of certain parameters to uncertainty, I am unsure of next steps. Could you discuss next steps?
20. Line 459: The sentence is unclear ("future work with subdivide")