

Review of "Assessing the performance of climate change simulation results from BESM-OA2.5 in comparison to a CMIP5 model ensemble" by V.B. Capristrano et al. (revised manuscript)

### Overall assessment and recommendation

This manuscript has been considerably improved in response to the first review process. In particular, use of English language has been thoroughly upgraded by external consultants and - as a consequence - reading to the text has become much smoother now. Given that the objective of GMD papers is proper description and evaluation of model tools rather than coming up with innovative scientific results, I find that the paper is now on a good way towards acceptance for this journal. As I already stated in the 1<sup>st</sup> review stage, I regard BESM to be a reasonable climate model that ought to get its chance for being used for specific scientific research.

Such credible progress notwithstanding, some more work on the manuscript is still necessary. First, there are still some (though not many) inconsistencies in the results presentation that require revision. In some cases this might be caused by excessively technical language that does not always help to address scientific interrelations precisely. Second, the language upgrade (though successful in an overall sense) has failed on occasions where the language editor apparently did not comprehend the scientific meaning of some text. The last three paragraphs of the conclusion section form a prototype of what I mean: The penultimate paragraph (p. 13, l. 25f) reads perfectly convincing and clear, whereas the two paragraphs surrounding it (p. 13, l. 15f and p. 14, l. 1f) are still very hard to grasp. Language polishing thus forms the bulk of my remaining recommendations and correction requests. This time, however, it will be sufficient if the authors carefully consider the list of proposed changes.

### General recommendations

- 1) p. 1 l. 3: I recommend to extend to "... precipitation, atmospheric circulation, and radiative feedbacks", given the room devoted to feedback analysis.
- 2) p. 2 l. 28: "... uncertainty in estimates of total precipitation due to uncertainties in ..."; Could it be preferable to replace "uncertainty" by "inter-model deviation"?
- 3) p. 6 Eq. 4: Despite the statements made in your reply I am still at a loss to detect your equation (4) in any of the given references. Particularly the splitting with respect to  $T_s$  and  $T_{as}$ .
- 4) p. 6 l. 24: "... they are shifted to the residuum." Concerning stratospheric temperature and the corresponding rapid radiative adjustment, this must make up for a quite large residuum then (see Smith et al., 2018, their Fig. 3). However, as you are not using the kernel derived feedbacks for closing the forcing vs. feedback balance, I will not insist on this point any further.

- 5) p. 7 l. 2: "This approach, however, assumes ..." – I do not understand what is implied by this sentence. Does the following proposed text meet your intention? "However, the resulting kernel derived feedbacks can only be assumed to reflect the actual feedback in the considered models under the premise of small differences between the radiative transfer codes. This has not been verified here, thus enabling inconsistencies between feedbacks derived from the regression method and the kernel method. On the other hand, Figure 3 indeed suggests small differences at least arising from use of either the NCAR or the GFDL radiative transfer model."
- 6) p. 8 l. 12: "Thus the ratio of ..."; as it sounds somewhat trivial that  $1/\lambda$  (the "climate sensitivity parameter") explains ECS (the "equilibrium climate sensitivity") well, you might consider a change to: "Thus inter-model variations in the balance of feedbacks explains the dispersion in the ECS better than ..."
- 7) p. 8 l. 22: "... the inter-model differences are greater than the distribution of the radiatively active constituents of the base model"; This sentence is rather confusing: Which differences? Which constituents? I suspect that you intend to address the issue of point 5 above, but it's not very clear.
- 8) P. 9 l. 2: "... that BESM does not have a higher contrast between the surface and upper troposphere temperatures in comparison to temperature contrasts of the other models."; a still quite clumsy sentence and also counter-intuitive, as in case BESM shows a more positive lapse rate feedback than the ensemble, this should imply that the warming is relatively larger at the surface and relatively weaker at the upper troposphere, resulting in an in fact more strongly increased (!) vertical temperature gradient. That seems to contradict your statement. Do you agree? Please, reformulate accordingly.
- 9) **Major point!** The interpretations starting at p. 10, l. 4, with respect to Fig. 6, still lack clarity and also inherent consistency. First, it is not obvious how the term 'cloud mask' is to interpret physically, as it is not mentioned when the 'adjusted cloud radiative effect',  $\Delta CRE_a$ , is introduced by Eq. (5). As a consequence, the link between the analytic explanation in the penultimate paragraph of 4.2 and the discussion of physical parameters in the last paragraph remains obscure. In particular (my second main point) it is not clear to me how (l. 15) "Because of the increase in total cloud fraction, a *negative* SW CRE appears ..." can be reconciled with the SW cloud feedback in Fig. 5 (mid right panel) being obviously *positive* in those very regions. Does the lower troposphere cloud response in BESM compete with the sea ice retreat differently than in the CMIP ensemble? While a full explanation is not required in the framework of a GMD paper, still you should try to explain the parameter inter-relations in a consistent and lucid way. Especially, because the concluding section resumes this point to emphasize it as a main part of the paper.
- 10) p. 13 l. 20: "... increase in cloud cover ... which increased the outgoing SW radiation at the TOA." ; see previous point 9! In Fig. 5 the SW cloud feedback in that latitude range is positive, which suggests that more SW energy remains in the climate system, thus indicating *less reflection*. By the way, "outgoing SW radiation" ought to be replaced by "reflected solar radiation".

- 11) p. 13 l. 3: " ... overcome by the albedo feedback cloud mask ..." ; also related to major point 9. This explanation is not suggestive and not adequately pointed out in a physical sense.
- 12) p. 14 l. 2: "In this sense it has contributed ..." You obviously refer to that part of the manuscript I have addressed in my general points 9-11. Evidently, I do not agree to this claim as not convincing argument for a "better understanding" has been provided. Nor do I demand or expect, that such a contribution to better understanding is given in the framework of a GMD paper. Hence, the statement could simply be discarded, or modified to something like "... the analysis methods used here have the potential to explain remaining process uncertainties causing inter-model spread in the cloud feedback in future work."

### Language and Technical Remarks

- 1) p. 1, l. 7 (Abstract): " ... the BESM simulations yield ..."
  - 2) p. 3 l. 10: "... physical processes ..." → "... physical parameterisations ..." ;  
"... those used by Veiga ..." → "... those discussed in Veiga ..."
  - 3) p. 3 l. 12: " ... is used in this study but with a ..." → "... is generally used in BESM-OA2.5, except for a ..."
  - 4) p. 3 l. 29: "... the value observed by the RAPID project ..." → "... the value determined within the RAPID project ..."
  - 5) p. 4 l. 1 (heading): "Comparison to a previous version" → "Comparison to a previous model version"
  - 6) p. 4 l. 2: "evolution" → "advancement"
  - 7) p. 4 l. 5: " ..., which results in ... to approximately  $-4 \text{ Wm}^{-2}$  ..." → "..., with a reduced global mean bias of approximately  $-4 \text{ Wm}^{-2}$  ..."
  - 8) p. 4 l. 26: "... presented less consistent results" → "yielded less consistent results"
  - 9) p. 5 l. 1: "... both experiments were run in parallel for 150 years." → "... for years 151 to 300 both scenarios are run in parallel to the piControl reference for 150 years." [I understand that for the evaluation of piControl, e.g. means of the reference state, only the last 150 years of piControl have been used. If, however, reference means are averaged over 300 years, this should be mentioned explicitly in the text.]
  - 10) p. 5 l. 7: "... two different methods: regression (Gregory et al., 2004) and radiative kernel ..." → "two different methods, using either a regression according to Gregory et al. (2004) or radiative kernels ..."
  - 11) p. 5 l. 14: "... method consists of the ..." → "... method consists in the ..." [or "... method involves the ..."]
  - 12) p. 5 l. 15: " ... the net radiation change ..." → "the net radiative flux change ..."
  - 13) p. 6 l. 1 (heading): "Climate feedbacks (radiative kernel)" → "Separating individual climate feedbacks using radiative kernels"

- 14) p. 6 l. 7: "... kernels consist of the impact in the radiative balance in the TOA via arbitrary increases in the ..." "... kernels represent the impact on the radiative balance at TOA via arbitrary increases of the ..."
- 15) p. 6 l. 8: "For *calculating* the temperature kernel, an *increment* of 1 K is added ..."
- 16) p. 6 l. 9: "For the albedo kernel ..."
- 17) p. 7 l. 1: "the models analyzed" → "the CMIP model analyzed"
- 18) p. 7 l. 23: "These linear regressions ..."; I recommend the following formulation of this sentence as an improvement: "The linear regressions based on all-sky radiative fluxes are used to estimate ECS,  $G$  and  $\lambda$ , while the regressions based on clear-sky data are used to obtain  $\Delta CRE$  ..."
- 19) p. 8 l. 2: "... were assessed as previously performed by Andrews et al. (2012) for 15 ..." → "were compiled in extension of the previous work by Andrews et al. (2012), who evaluated 15 ..."
- 20) p. 8, l. 4: "For the 15 models ..."; I recommend to reformulate to: "In Andrews et al. (2012) the ECS ranges from 2.07 to 4.74 K for the 15 models analyzed there, which is largely confirmed by our analysis."
- 21) p. 8 l. 5: "The possible small differences can be attributed ..." → "The small differences can possibly be attributed ..."
- 22) p. 8 l. 14: "... that is *well* within the range of ..." [certainly an argument in favor of your BESM].
- 23) p. 8 l. 25: "The Planck feedback global-mean is negative ..." → "The global mean Planck feedback is [strongly] (?) negative ..."
- 24) p. 8 l. 28: "... which is an overestimation compared to the ensemble mean value" → "... slightly underestimating the ensemble mean value in magnitude"; Technically, your formulation is of course correct, but "overestimation" is somewhat counter-intuitive to suggest that it is in fact "less negative".
- 25) p. 8 l. 31: I recommend adding the following introductory sentence: "Figure 4 shows the latitudinal profiles basic to the global mean feedback values of Figure 4, allowing to identify the regions that induce deviations of BESM results from the CMIP ensemble. In Figure 4a-b, there is ..."
- 26) p. 8 l. 33: "... with nearly the same increased ..." → "... with a similarly increased ..."
- 27) p. 8 l. 34: "... one of the lowest values ..." → "... one of the most negative values ..." [see point 24 above]
- 28) p. 9 l. 1: "(with respect to the surface temperature)" → "(corresponding to large surface warming)"
- 29) p. 9 l. 5: "adiabate"
- 30) p. 9 l. 6: "... is greater at the upper troposphere ..." → "... is larger in the upper troposphere ..."

- 31) p. 9 l. 8: " ... is responsible for a potentialization of the greenhouse gas effect, revealing a positive ..." → "... is causing a reinforcement of the greenhouse effect, reflected by a positive ..."
- 32) p. 9 l. 9: " ... of this *close* link ..."
- 33) p. 9 l. 10: " ... it is common to sum their effect, as performed in ..." → "... it is common to consider their effects as the sum, as displayed in ..."
- 34) p. 9 l. 11: "The greatest BESM deviations ... for all models"; Does the message of this statement differ from what has been told on p. 9, l. 2-4 ? [ see also general remark No. 8) If not, my recommendation is to discard this sentence.
- 35) p. 9 l. 16: "The albedo feedback values computed for BESM and other CMIP models are show in ..." "The albedo feedback profiles from BESM and the CMIP models are compared in ..."
- 36) p. 9 l. 16: "These results are particularly important over ..."; somewhat fuzzy, better change to "Non-zero results mostly occur over ..."; You might also consider to support this statement by citing evidence from papers that have presented corresponding geographical distributions like Chung and Soden (2015, their Fig. 9) or Block et al. (2020, their Fig. 2a). Accordingly, respective information for the Planck feedback (p. 8 l. 32) is provided by Rieger et al. (2017, their Fig. 2).
- 37) p. 9 l. 18: " ... all of the models ..." → "all models"
- 38) p. 9 l. 20: " ... how fast the sea-ice melts in ..."
- 39) p. 9 l. 21: "showed" → "show"
- 40) p. 9 l. 23: "Pithan and Mauritsen"; here, too, the recent paper from Block et al. (2020) [see point 36] could be a very appropriate reference.
- 41) p. 9 l. 31: "... as they reported a nearly neutral and positive ..." → "..., who also reported a near-zero to positive ..."
- 42) p. 9 l. 35: "... in the high latitudes." → "... at high latitudes."
- 43) p. 10 l. 2: " ... assess the *analytical* causes ..." [to distinguish from the more physical cause addressed in the subsequent paragraph]
- 44) p. 10 l. 5: "... all-minus clear-sky radiation ..." "...all-sky minus clear-sky radiative flux ..."; However, the meaning of the sentence is hard to comprehend anyway; see general comment No. 9.
- 45) p. 10 l. 6: "We observe that ..." → "We find that"
- 46) p. 10 l. 9: "... it is possible to see that the major contributor to BESM's status ... is the SW CRE" → "... it is possible to attribute BESM's status ... to the SW CRE ..."
- 47) p. 10 l. 12: "A further analysis ..." → "A deepened physical analysis ..."
- 48) p. 10 l. 27: " ... in the air temperature ..." ??? - Do you mean "... in the surface temperature ...", which is what is shown in Figure 8 ?
- 49) p. 10 l. 29: "shows" → "yields" (2x)

- 50) p. 10 l. 32: "... have similar ..." → "... have closely similar ..."; "... there are few spatial correspondences between ..." → "... there is hardly any pattern correlation between ..."
- 51) p. 11 l. 4: "..., which show ..." → "..., who show ..."
- 52) p. 11 l. 7: "... mainly over the North Pole." → "... particularly over the Arctic" [certainly the North Pole as a specific point in space does not show maximum warming].
- 53) p. 11 l. 8: "... shows an increase in temperature ..." → "... shows a relative maximum in warming ..." [temperature increase occurs almost everywhere!]
- 54) p. 11 l. 14: "... not possible to note a linear ..." → "... not possible to claim any robust linear ..."
- 55) p. 11 l. 17: "Plank" → "Planck"
- 56) p. 11 l. 20: "... are approximately similar to ..." → "... are, in general, similar to ..."
- 57) p. 11 l. 25: "... notable discrepancies." → "... notable exceptions."
- 58) p. 11 l. 31: "... suggests a linear ..." → "... emphasizes a linear ..."; "indicating indicating"
- 59) p. 11 l. 33: "As shown in ..." → "As obvious from Figure 9, the BESM performance perfectly matches the ensemble mean behavior in the global mean."
- 60) p. 12 l. 4: "The apparent ..." → "An apparent ..."
- 61) p. 12 l. 19: "This SLP decrease ...indicate a ..." → "This, connected with the increase in the mid-latitudes, indicates a ..."
- 62) p. 12 l. 32: "Figure 13 shows ..." ; this sentence needs rectification. First, I notice no "zonal wind" in the figure. Second, your explanation in the reply notwithstanding, the difference between "omega" and "omega vertical motion" is not clear. You might consider to introduce the term of "generalized vertical velocity" (omega [pa/s]) to distinguish it from the physical vertical motion (w [m/s]). If, however, such a distinction is irrelevant for the interpretation of the figure, you might for simplicity just replace "omega" by "vertical velocity".
- 63) p. 13 l. 2: "... positive values ..." → "... enhanced subsidence ..."
- 64) p. 13 l. 7: "... climate sensitivity parameters ..." → "... key sensitivity parameters ..."
- 65) p. 13 l. 16: "... considerable standard deviations ..." → "... considerable inter-model variability ..."
- 66) p. 13 l. 17: "... BESM also showed cloud feedback values larger than the zonal mean plus standard deviation for the analyzed models ..." ; that's not easy to comprehend, do you mean "... the BESM zonal mean cloud feedback ranges outside the standard deviation for the analyzed models ..." ?
- 67) p. 13 l. 19: "... was approximately  $0 \text{ Wm}^{-2}\text{K}^{-1}$ ." → "... was close to zero."
- 68) p. 13 l. 32: "... was clearer in ..." → "... was more distinct in ..."

- 69) p. 14 l. 1: "BESM is ..." "Summarizing, we conclude that BESM-OA2.5 is a climate model that can reproduce approved physical processes that determine and modify changes of the global climate system."
- 70) p. 14 l. 4: " ... the BESM development team is committed to improving the cloud cover of the model as well as ..." → "... the BESM team continues its effort to improve the cloud parameterization of the model as well as ..."
- 71) p. 14 l. 5: " ... is under revision." That could easily used as an argument to reject the paper. Hence, I recommend " ... is seen as an issue that is to be tackled in ongoing model development work."
- 72) p. 14 l. 6: "... it will be more compatible ... ocean an atmosphere." → "... it will include physical parameterizations of atmosphere/ocean interaction that lead to better agreement with other model and with observations."
- 73) Figure 5: Are these feedbacks as the figure caption suggests? Then the label of the colour bar should be " $\text{Wm}^{-2}\text{K}^{-1}$ ".
- 74) Figure 6: There is no indication of physical units in that figure or figure caption.
- 75) Figure 7: There is no indication of physical units in that figure or figure caption.
- 76) Figure 8, caption: "...isoline..." → "contour line"
- 77) Figure 9, caption: "...isoline..." → "contour line"
- 78) Table 3, caption:  $\Delta\text{CRE}$  ought to have unit  $\text{Wm}^{-2}$ .

#### References (only if not already cited in the paper):

- Block, K. et al., 2020: Climate models disagree on the sign of total radiative feedback in the Arctic, *Tellus A*, 72, 1696139.
- Chung, E.-S. and Soden B.J., 2015: An assessment of direct radiative forcing, radiative adjustments, and radiative feedbacks in coupled ocean-atmosphere models, *J. Clim.*, 28., 4152-4170.
- Rieger, V.S. et al., 2017: Can feedback analysis be used to uncover the physical origin of climate sensitivity and efficacy differences? *Clim. Dyn.*, 49, 2831-2844.
- Smith, C.J. et al., 2018: Understanding rapid adjustments to diverse forcing agents, *Geoph. Res. Lett.* 45, 12023-12031.