# Response to reviewer 1

*General Comments:*

*Sanderson et al. use a simple climate model composed of several exponential decay terms to model the output of pre-industrial control and abrupt-4×$CO_2$ simulations from CMIP5, CMIP6, and LongRunMIP. The authors use this simple climate model to estimate potential biases in effective climate sensitivity (EffCS) estimates. This approach is novel and provides an interesting framework to analyze EffCS; however, there are several points that the authors should address in order for me to recommend this manuscript for publication.*

Many thanks for the careful review and recommendations.

*Specific Comments:*

*1) The new framework the authors developed is interesting; however, I am having a hard time deciphering why this paper is important. I am not sure what the main point of the paper is. Is the main point to answer the question the authors stated at the end of the introduction: "How plausible are the higher sensitivity [CMIP6] models"? Is the main point to say that ECS is actually higher than suggested by EffCS given by CMIP6 or IPCC AR6? The authors state "Our results highlight the potential for error in estimates of effective climate sensitivity through the assumptions on the asymptotic radiative balance of climate models (page 9 line 9)". The authors need to go a step further and provide an indication of what their suggestion for the value of EffCS would be based on their new framework. The authors should discuss their results in the context of recent literature that examines estimates of EffCS. Recent studies have provided estimates of EffCS, such as Zelinka et al. (2020), Tokarska et al. (2020), McBride et al. (2021), Sherwood et al. (2020), and the new comprehensive evaluation conducted by IPCC AR6.*

*Do the authors have a new range of EffCS using their approach compared to these other analyses? Could the authors suggest a way to constrain the estimate of EffCS based on the model's radiative imbalance between the PICTRL and ABRUPT4X simulation? The authors should add comparisons to recent literature in their results section. In the conclusions section, the authors should expand upon the importance of their results to indicate a revision or addition to current estimates of EffCS, or suggestions on how to revise the current estimate of EffCS using their approach.*

Thanks for this point. We agree that the introduction perhaps could better clarify the objectives of the paper. The study does not seek to provide a new estimate of real-world climate sensitivity. This is a somewhat orthogonal issue to the central finding in the study that estimates of model climate sensitivity based on the Gregory regression approach are likely to be biased by different stable energetic states under pre-industrial carbon dioxide concentrations and quadrupled CO2 concentrations. On revision, we will clarify in the introduction that this study is a comment on the standard approach for calculating effective climate sensitivity, and not on the likelihood of any given value.

We do, however, now highlight that some aspects of the 'hot model' problem may be resolved by the issues raised in this study - in that the model current estimates of Effective Climate Sensitivity are in many cases not reliable. Quantifying this bias in CMIP6 is not possible with existing simulations, but for LongrunMIP - this bias results in a consistent over-estimate of EffCS in 5 of 15 models (with no examples of the opposite).

We expect that the corrections implied by this study will have some impacts on our assessment of likely ranges of ECS - especially for emergent constraint studies which have identified relationships between published ECS values and observable quantities. On revision, we will consider to what degree existing literature (including the AR6 assessment) could be impacted by the findings of this study. A revised assessment of ECS is, however, out of scope for this study.

*2) How did the authors determine the minimum and maximum values of τ for the short timescale, intermediate time scale, and long-time scale given in Table 1? Are these values supported by literature?*

Our priors on timescales are informed by previous studies considering a 3 timescale decomposition of Earth system transient response to abrupt 4xCO2 forcing (e.g. Proistosescu, C., & Huybers, P. J. (2017), Rugenstein and Armour 2021, Caldeira and Myhrvold (2013)), noting that the FaIR model used in the IPCC AR6 assessment allows for only 2 timescales (Smith 2018). There is no universal approach in these former studies for priors on the different timescales. Given this, we here choose relatively uninformative priors compared to previous studies, broadly in place to order to structure the response modes into subdecadal, decadal, centennial, millennial and multi-millennial modes without being overly prescriptive. On revision, we will clarify our logic on this point. We also now consider a sensitivity study of how allowance of different timescales in the model impacts our results.

*3) The authors should explain how assessing the radiative imbalance in the control simulation , $R_{0CTRL}$, impacts the parameters in Eq. 1a or Eq. 1b. As currently written, it is unclear how this assessment is incorporated into Eq. 1a and 1b.*

The control simulation radiative imbalance does not directly impact equations 1a,b. It is assessed independently from the control simulation and compared with the asymptotic energetic imbalance in the abrupt4x simulation.  We have now formulated $R4x_{extrap}$ (the equilibrium TOA balance state following the abrupt4x perturbation) to be a parameter in the fitted equation as suggested.  We agree that this is a more accurate way of presenting our approach (it makes no difference to the actual derived fits, which were effectively always derived using $R4x_{extrap}$).

45
$$T(t) = \sum_{n=1}^{3} S_n(1 - e^{-(t/\tau_n)}) + T_0$$

$$R(t) = \sum_{n=1}^{3} R_n(-e^{-(t/\tau_n)}) + R^{4x}_{extrap}$$

*4) Equilibrium climate sensitivity and effective climate sensitivity are the response of the climate system to a doubling of $CO_2$ relative to preindustrial. The authors use the ABRUPT4X scenario, which is for a quadrupling of $CO_2$. In other methods, such as Gregory et al. (2004), the temperature response to the quadrupling of $CO_2$ needs to be divided by 2 to achieve an estimate of the temperature response to the doubling of $CO_2$. The authors do not discuss how their method accounts for the fact they are using an ABRUPT4X scenario to assess the temperature response to a doubling of $CO_2$. The authors should elaborate in the methods section how they account for this discrepancy.*

We have added the following to explain our policy on reporting equilibrium climate sensitivity:

*We report climate sensitivities for a doubling of carbon dioxide from pre-industrial levels. As such, we follow standard practice in dividing ABRUPT4X sensitivities (EffCS, $\Delta$Textrap and $\Delta$Tbest−est) by 2 (Meehl et al., 2020), though we note that in some models this approximation introduces minor errors (Jonko et al. (2012); Rugenstein et al. (2019)), this is not the focus of the present study.*

*5) There is no mention of IPCC AR6 in this paper. How does this analysis compare to the best estimate (3°C) and range of (2 - 5°C) of ECS given by AR6? Does the new framework in this paper support a lower or higher value of EffCS than provided by IPCC?*

Our paper does not contain an estimate of real-world climate sensitivity, rather it is a comment that existing estimates of model climate sensitivity may be unreliable.  This does however have implications for the use of Effective Climate Sensitivity to assess the reliability of a given model. We now discuss this both at the end of the introduction, and in the conclusions, with explicit reference to AR6 (which was not final when this paper was first drafted).

We have also added a longer discussion on the implications of this study for future assessment:

> *This directly impacts our ability to accurately measure $EffCS$ from short simulations, and draws into question whether $EffCS$ should be used as a factor at all in assessing the fidelity of climate models (Hausfather, 2020). Effective climate sensitivity has known limitations that it describes effective feedbacks at a certain representative timescale following a change in forcing (Rugenstein 2021), but our results here highlight another issue that EffCS can only be used if we can be confident in the asymptotic energetic balance of the model. Such confidence can arise either from a ground-up demonstration of structural energy conservation in the model (Hobbs 2016), or by running sufficiently long simulations to be empirically confident both in the pre-industrial energetic balance and in the asymptotic multi-millennial tendencies of the model following a change in climate forcing. However, such experiments are currently difficult to achieve for CMIP class models, the 5000 year simulations conducted in Rugenstein (2020) were significantly longer than any experiments conducted previously - and we find in the present study that they remain too short to have confidence in the asymptotic state.*
>
> *Given this, this study has multiple recommendations. Firstly, a greater emphasis in climate model design and quality checking needs to be placed on structural closure of the energy budget in the climate system. Models which can demonstrate that energy is conserved in the model equations can allow confidence that the system as a whole will converge to a state of true radiative equilibrium following a perturbation, which would allow a robust calculation of $EffCS$. For models which cannot demonstrate this, longer simulations are required to be confident in the asymptotic state - but these simulations may be prohibitively time and resource consuming. Such limits could potentially be alleviated through the use of lower resolution configurations (Kuhlbrodt 2018, Shields 2012) (with the risk that such models will exhibit different feedbacks from their high resolution counterparts) or by considering analytical approaches to accelerate convergence of complex systems (Xia 2012) . However, in the short term a more practical approach may be to consider alternative climate metrics which do not require assumptions about the equilibrium state of the system. Transient Climate Response does not require assumptions about radiative flux, but it does not provide direct information on the warming expected under stabilising forcing. A possible alternative is A140 (the warming observed 140 years after a step quadrupling in CO2 concentrations (Sanderson 2020), which requires no assumption on equilibrated state - and is more informative on the warming expected under high mitigation scenarios than $EffCS$ itself (even if it is known without bias due to energetic leaks). In conclusion, the role of Effective Climate Sensitivity as a metric in assessing the response of the climate system should be reconsidered, both due to its lack of relevance to projected warming under mitigation scenarios (Knutti 2017, Frame 2006, Sanderson 2020) but also due to the fact that its derivation requires assumptions about the asymptotic state of the climate system which cannot be demonstrated in a number of Earth System Models.*

*6) Figures 1 and 2 are barely discussed. The authors should add more discussion of these figures to the results section, especially highlighting any important interpretations of the figures, or move these two figures to the Appendix.*

We consider these figures to be a key part of the discussion, but there were insufficient references. A number of explicit figure references have been added where these Figures are relevant.

*7) In the results section, the authors jump back and forth between discussing Figure 3 or Figure 4 (Page 6 lines 1 – 19), making it difficult to follow the points the authors are trying to make. The authors should consider editing this section by first discussing and interpreting Figure 3, then discussing and interpreting Figure 4.*

Thanks - this text has been reformatted

*8) The authors need to verify that the figure captions match the figures. Colors and types of lines described in the figure captions do not match what was plotted in the figure, making it difficult to interpret the figures (see the Technical Corrections related to each figure below).*

Figures have been extensively revised following reviews and clarified throughout. Apologies for caption inaccuracies in the submitted version.

*9) Table A2 is an important table, displaying the difference between EffCS computed using various methods for the LongRunMIP simulations. The authors should consider moving Table A2 into the main part of the text. They can add a discussion of the table to the results section, highlighting why the estimates for $\Delta T_{best\text{-}est}$ and $\Delta T_{extrap}$ are similar for some models yet different for others.*

Table moved to main text, thanks for the suggestion.

*Technical Corrections:*

*Equation 1b: Constant is written as $R_{4x}$, but referred to as $R_{04x}$ in the text (page 3 line 1)*

Revised, thanks.

*Table 1: $R_n$ scaling factors are not listed in Table 1, but $S_n$ scaling factors are listed. Is there a reason why the $R_n$ scaling factors are omitted?*

Now included, thanks

*Table 1: $R_0$ is included in the table, but this variable does not appear in either Eq. 1a or Eq. 1b. How does this variable relate to these two equations?*

A typo using notation from a former version. Corrected, thanks.

*Why are the lines in figures 1 and 2 labeled as $S_{LR}$, $S_{eff}$, and $S_{extrap}$. In Eq. 1a, 1b, 2, and 3, S refers to a scaling factor. Why are the authors using this variable (S) to label the different lines?*

Notation from a former version. Corrected, thanks.

*Figure 1 Caption:*

- *Authors state solid yellow lines are linear regressions used to estimate effective climate sensitivity for the first 150 years of data. This should be the dotted yellow lines.*

Thanks, corrected.

- *Authors state solid pink lines are linear regressions used to estimate effective climate sensitivity for the last 15% of warming. This should be the dotted pink lines.*

Thanks, corrected.

- *Authors state vertical dotted pink and yellow lines show corresponding values of effective climate sensitivity. Should be vertical solid pink and yellow lines.*

Thanks, corrected.

- *Authors state solid yellow horizontal line shows the PICTRL net energy imbalance averaged over the final 100 years of the simulation. There are no solid yellow horizontal lines. There are green horizontal lines, which are not included in the caption or legend. Are the green lines supposed to be the PICTRL net energy imbalance? If not, make sure to label what the green lines are showing.*

  Legend added.

- *Solid blue line is not described in the caption*

  Added.

- *I am not sure that the dashed blue line is described correctly in the figure caption. Authors say the dashed blue line shows an exponential model fit, but the lines in all of the subplots in Figure 1 are horizontal. Is the solid blue line actually showing the exponential model fit? If so, what do the dashed blue lines represent?*

  Sentence rewritten, legend entry added.

- *Green dots are not described in the caption*

  added

*Figure 1 General Comments:*

- *Green and blue dots in the legend representing PICTRL and ABRUPT4X are very faint, almost impossible to see. Make them more legible in legend.*

Removed from legend (high transparency points do not show well in legend, but are necessary to see the shape of the point distribution). Description remains in caption.

- *I cannot distinguish the difference between the blue dots representing ABRUPT4X and the light blue ellipse showing the 5-95 CI for $\Delta T_{extrap}$. It looks like only the light blue ellipse is plotted.*

In the longrunmip fits, the 5-95 CI is negligible - the ellipse is only visible on the CMIP5/CMIP6 fits in the supplementary. We've removed the legend entry to avoid confusion.

- *What does $n_{yr}$ show? I assume it is the number of years in the LongRunMIP simulation, but the authors should include a description of the parameter in the figure caption for clarity.*

Text removed

- *Make sure the lines plotted on the figure do not go through the text (i.e., CNRMCM61 panel has solid blue and dotted yellow lines going through $n_{yr}$ = 1850)*

Text removed.

*Figure 2 Caption:*

- *There is a description of black points, but there are no black points in the figure or legend*

Corrected

- *Which dashed horizonal line illustrates $\Delta T_{extrap}$? Blue? Green?*

Corrected

- *A description of the green dashed line does not appear in the figure caption, and the green dashed line is not included in the legend.*

Corrected

- *A description of the green dots does not appear in the figure caption*

Corrected

- *Authors state the dashed purple line is $\Delta T_{best\text{-}est}$. I do not see a purple line. There is solid pink line. Is this pink line supposed to be $\Delta T_{best\text{-}est}$?*

Corrected. Yes, pink line.

*Figure 2 General Comments:*

- *Missing "of" in the sentence: "Shaded regions and thin dotted lines show the 10th and 90th percentiles of the fitted ensemble projections"*

  Added.
- *The $4xCO_2$ and pictrl is written differently from PICTRL and ABRUPT4X in the first figure caption and the main text. These scenarios should be referred to in a consistent manner*

Corrected
- *There are no lines or symbols next to $4xCO_2$ and pictrl in the legend*

Removed (description now in caption).

*Figure 3 General Comments:*

- *It is difficult to distinguish the blue dots and the blue shaded region, specially towards the right side of each panel. Making the shaded region a different color, or different shade of blue could help distinguish the points from the shaded region.*

  Point Color changed to grey
- *Why do some of the models have visible 10th and 90th percentiles at the beginning and ending of the blue line, but others do not? What is different in the models with very small ranges of uncertainty from those with larger ranges?*

Shorter simulations (of *only* 1000 years) are subject to larger uncertainties in the fit for timescales >1000 years. This is evident for MPIESM12 and ECHAM5MPIOM. This feature is now noted in the first paragraph of the results.
- *$4xCO_2$ in the legend does not match ABRUPT4X labeling in figure caption and the main text*

corrected.
- *Missing "of" in the sentence: "Shaded regions and thin lines show the 10th and 90th percentiles of the fitted ensemble projections"*

corrected.

*Figure 4 Caption:*

- *Left hand column:*
  - *Caption says there are whiskers in the left-hand column on the light blue diamond symbols. There are no whiskers plotted showing the 10th & 90th percentiles of $\Delta T_{extrap}$*

corrected

- *Central Column:*
    - *Caption says there are cyan error bars plotted, but they are not on the figure. Only show blue diamonds*

corrected

- *Solid and dashed yellow lines are not described in the figure caption*

Subfigure deleted

*Figure 4 General Comments:*

- *Is there any range of uncertainty for the values of $\Delta T_{best\text{-}est}$ shown by the red diamonds? If so, then this uncertainty should be indicated on the figure*

No, not as described in Rugenstein 2020.

- *There is no legend included with this figure, whereas the other 3 figures included legends. Consider adding a legend to this figure.*

Done

*Page 3 Line 31: What does "this estimate" refer to? $\Delta T_{extrap}$, $R_{extrap4x}$, or both?*

Both - now clarified in text

*Page 3 Line 35: Some other models should be included as described as behaving as expected. GISSE2R and GFDLESM2M show near zero equilibrium TOA balance in both PICTRL and ABRUPT4X simulation in Figure 3. Why were these models excluded from this sentence?*

Added, thanks.

*Why are the values in the brackets for $\Delta T_{extrap}$ and $\zeta_{extratp}$ the same in Tables A2, A3, and A4? The table caption explains that the numbers in the brackets represent the $5^{th}$ and $95^{th}$ percentiles. I find it highly unlikely that the $5^{th}$ and $95^{th}$ percentiles are the same, especially since the median value is larger than the values in the brackets.*

There was an typo in the code creating this table - 95th percentile is now correctly printed

*Page 9 Line 21: Missing closing parentheses after Table A1*

Fixed, thanks.

**References**

Hausfather, Zeke, Henri F. Drake, Tristan Abbott, and Gavin A. Schmidt. "Evaluating the performance of past climate model projections." Geophysical Research Letters 47, no. 1 (2020): e2019GL085378.

Proistosescu, C., & Huybers, P. J. (2017). Slow climate mode reconciles historical and model-based estimates of climate sensitivity. Science advances, 3(7), e1602821.

Rugenstein, Maria AA, and Kyle C. Armour. "Three flavors of radiative feedbacks and their implications for estimating equilibrium climate sensitivity." Geophysical Research Letters 48, no. 15 (2021): e2021GL092983.

Caldeira, K., and N. P. Myhrvold. "Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration." Environmental Research Letters 8, no. 3 (2013): 034039.

Smith, C. J., Forster, P. M., Allen, M., Leach, N., Millar, R. J., Passerello, G. A., & Regayre, L. A. (2018). FAIR v1. 3: a simple emissions-based impulse response and carbon cycle model. Geoscientific Model Development, 11(6), 2273-2297.

Meehl, Gerald A., Catherine A. Senior, Veronika Eyring, Gregory Flato, Jean-Francois Lamarque, Ronald J. Stouffer, Karl E. Taylor, and Manuel Schlund. "Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models." Science Advances 6, no. 26 (2020): eaba1981.

Jonko, Alexandra K., Karen M. Shell, Benjamin M. Sanderson, and Gokhan Danabasoglu. "Climate feedbacks in CCSM3 under changing CO 2 forcing. Part II: variation of climate feedbacks and sensitivity with forcing." Journal of Climate 26, no. 9 (2013): 2784-2795.

Rugenstein, Maria, Jonah Bloch-Johnson, Ayako Abe-Ouchi, Timothy Andrews, Urs Beyerle, Long Cao, Tarun Chadha et al. "LongRunMIP: motivation and design for a large collection of millennial-length AOGCM simulations." Bulletin of the American Meteorological Society 100, no. 12 (2019): 2551-2570.

Hobbs, Will, Matthew D. Palmer, and Didier Monselesan. "An energy conservation analysis of ocean drift in the CMIP5 global coupled models." Journal of Climate 29, no. 5 (2016): 1639-1653.

Kuhlbrodt, Till, Colin G. Jones, Alistair Sellar, Dave Storkey, Ed Blockley, Marc Stringer, Richard Hill et al. "The low‐resolution version of HadGEM3 GC3. 1: Development and evaluation for global climate." Journal of Advances in Modeling Earth Systems 10, no. 11 (2018): 2865-2888.

Shields, C.A., Bailey, D.A., Danabasoglu, G., Jochum, M., Kiehl, J.T., Levis, S. and Park, S., 2012. The low-resolution CCSM4. Journal of Climate, 25(12), pp.3993-4014.

Sanderson, B., 2020. Relating climate sensitivity indices to projection uncertainty. Earth System Dynamics, 11(3), pp.721-735.