

Response to reviewers' comments on egusphere-2026-163

Kyriaki Bimpiri, Thomas Hocking, Thorsten Mauritsen

Introduction

We would like to thank two anonymous reviewers for their comments. We are very grateful for their constructive feedback, which has helped us improve our manuscript. In the following document we provide a point-by-point response to their comments. Our responses are shown in blue.

HadCRUT5

In the initial manuscript, the calculations for Fig. 10 were accidentally made using the 5.0.2.0 version of the HadCRUT5 dataset, rather than the 5.1.0.0 version that was stated. The manuscript has been updated to use version 5.1.0.0, with only very minor changes in the calculations and the figure as a result, and no impact on the overall conclusions of the study.

Time period

We have extended our analysis with data from the year 2025. As a result, there are minor numerical changes to some of the figures, but the overall conclusions of the study remain the same.

Comments by Anonymous Referee #1

General comments

This study examines the recent positive trend and magnitude of the observed Earth's Energy Imbalance (EEI) and evaluates their reproducibility in CMIP6 models. The authors demonstrate that while the magnitude of EEI is well captured by CMIP6 models, its positive trend is underestimated by most of them. They further identify a close relationship between EEI and equilibrium climate sensitivity (ECS) and attempt to constrain ECS using the observed EEI magnitude, suggesting that models with moderate ECS values are more consistent with observations.

Overall, the manuscript is well written, and the results are interesting and supported by robust analyses. I recommend publication after minor revisions.

Specific comments

- L86–87:

I understood that a positive sign indicates energy input into the system, whereas a negative sign indicates energy loss. If this understanding is incorrect, please clarify. Based on Figure 2, the global mean temperature appears to increase when the energy imbalance is positive, which seems consistent with net energy input.

For a system with constant forcing that conserves energy you are absolutely right: a positive imbalance means that there is a gain of energy over time. However, climate models do not necessarily conserve energy [Mauritsen et al., 2012]. In a climate model that is run to a stationary pre-industrial control state, any leakage of energy in this control state needs to be compensated by a positive energy imbalance. Vice versa if the model artificially produces energy, the surplus must be emitted to space, corresponding to a negative energy imbalance.

We have updated the text to better explain this.

- L243–246:

I found this discussion somewhat confusing. You state that the cooler global temperature associated with internal variability during the first period leads to a larger energy imbalance through the negative feedback term ($\lambda \Delta T_s$). However, my understanding is that this term represents the response to anthropogenic forcing, while the effect of internal variability is captured by ϵ . Please clarify this point and elaborate on the underlying argument in more detail.

The ϵ term represents internal variability in the radiation balance that occurs independently of the global mean surface temperature, e.g. from random weather or unforced pattern effects. The underlying argument is that in the first period, a cooler global temperature than expected (i.e. a smaller ΔT_s) means that the magnitude of $\lambda \Delta T_s$ is smaller (i.e. the term is less negative) and hence that the imbalance is larger. We have clarified the description of the terms in Eq. 1 and elaborated on the underlying argument.

- L250–251:

I understand that models with larger ECS tend to have smaller (weaker) negative climate feedback parameters λ . However, the magnitude of $\lambda \Delta T_s$ also depends on the value of ΔT_s . My understanding is that larger-ECS models compensate for energy imbalance through a larger increase in ΔT_s compared to smaller-ECS models, suggesting that the behavior of $\lambda \Delta T_s$ may not be straightforward. Please clarify this interpretation.

Your understanding is correct. In the case of high-ECS models, a small feedback parameter λ means that there is a weak response to the original forcing F . ECS is calculated at equilibrium, and the

surface temperature increase depends only on the forcing F and the feedback parameter λ . Things become more complicated during transient warming: the evolution of the energy imbalance and ΔT_s also depend on the ocean heat uptake coefficient when the surface temperature change is not yet fully manifested. For small ECS, λ is large and negative and the imbalance will be small, whereas for a very large ECS the restoring feedback is weak and the imbalance larger, even if more warming occurs. This was first pointed out by Hansen et al. [1985]. This effect is included when we run the two-layer model shown as a curve in Figure 11.

We have also clarified this discussion in our manuscript.

- L289-298:

As discussed, the observed sea surface temperature exhibits a La Niña-like trend pattern, whereas most climate models simulate an El Niño-like trend pattern. As shown in, for example, Gregory and Andrews [2016] and Andrews et al. [2018], climate feedbacks under real-world conditions are expected to be more strongly negative than those estimated from abrupt-4xCO₂ or historical experiments. If the external forcing term F and ΔT_s are the same in the models and in the real world, the observed energy imbalance N would be expected to be smaller. However, this does not appear to be the case based on observations. Could this discrepancy arise because F is not fully represented in the models, or because ΔT_s is relatively small in the observations? I would appreciate further clarification and discussion of this apparent inconsistency.

It is correct that in the past, a La Niña like pattern was present in the Pacific, and that this was something models had difficulties simulating; a widely studied problem. The warming pattern has, however, changed and the Pacific warming is closer to uniform in recent decades. A strong pattern effect, as discussed by Gregory and Andrews [2016] and Andrews et al. [2018] associated with La Niña-like conditions in the Pacific Ocean, during the first time period (1980-2005) would dampen the imbalance, through increased low-level clouds reflecting more shortwave radiation back to space. In the second time period (2006-2025) this temperature trend is reversed, with a more uniform pattern found in the observations. Climate models running the historical simulations do not fully reproduce this strong surface temperature pattern in the first time period, and ongoing research by others suggests that they also have difficulties simulating the 2006-2025 pattern [e.g. Zhou et al., 2016]. This could partly explain the model overestimation of the magnitude of the imbalance during the first time period, alongside the underestimation in the second period.

To clarify this discussion, we have added Figure 12 and provided a more detailed explanation in our manuscript.

Technical corrections

- L25:

Shared Socioeconomic Pathway SSP2-4.5 → Shared Socioeconomic Pathway (SSP) 2-4.5

We have added parentheses the first time we introduce the "SSP" abbreviation.

- Figure 4:

The gray lines are too thin to be easily distinguishable. It would be helpful to slightly thicken the lines or use different colors. In addition, please include information about these lines in the figure legend. We have updated Figure 4 with thicker grey lines and the inclusion of the Multi-Model Mean time series of the Energy Imbalance as well as the corresponding linear regression line.

- Figure 8:

"Multi-model ensemble mean/range" would be more appropriate than "Model mean/range."

Figure 8 has been adjusted.

- Figures 10 and 11:

The legends are too small. I suggest placing the legends outside the figures, increasing the font size, and arranging them in two columns.

The layout of these figures has been updated.

Comments by Anonymous Referee #2

"Climate models with moderate climate sensitivity best simulate the magnitude of Earth's energy imbalance" by Bimpiri et al. analyze coupled climate model simulations of Earth's energy imbalance and find that those with moderate ECS agree more with observations of EEI (magnitude). The paper is interesting and well-written. I only have minor comments below, mostly related to clarifications, wording, and references, that should be addressed prior to publication.

1. Para 15: Didn't Fourier and Arrhenius only describe the greenhouse effect? Not EEI? Please clarify.

Arrhenius in particular is perhaps more well known for his work on the greenhouse effect, but in fact both Fourier and Arrhenius describe the EEI, if not with that exact name:

"In fact, the radiation of the sun in which this planet is incessantly plunged, penetrates the air, the earth, and the waters; its elements are divided, change in direction every way, and, penetrating the mass of the globe, would raise its mean temperature more and more, if the heat acquired were not exactly balanced by that which escapes in rays from all points of the surface and expands through the sky." [Fourier, 1822, 1878]

"[...] there prevails an equilibrium in the temperature of the earth and of its atmosphere. The atmosphere must, therefore, radiate as much heat to space as it gains partly through the absorption of the sun's rays, partly through the radiation from the hotter surface of the earth and by means of ascending currents of air heated by contact with the ground. On the other hand, the earth loses just as much heat by radiation to space and to the atmosphere as it gains by absorption of the sun's rays." [Arrhenius, 1896]

2. Para 20: Underestimate, not "unable to", would be more appropriate. Please consider.
Done.

3. Para 25: von Schuckmann et al., 2016 would be an appropriate reference here. Please add.
This reference has been now added in our manuscript.

4. Para 40: Park and Soden, 2025 suggest a minimal aerosol role and should be mentioned. Please add.
We have incorporated this reference to our manuscript.

5. Para 70: The use of DEEP-C is concerning. It uses ERA5 which has a near-zero EEI trend due to faulty SWCRE changes (Loeb et al., 2022; Raghuraman et al., 2023). This should be acknowledged with these references (all references below). Please add.
This is a reasonable concern. As far as we can tell, the two suggested references do not explicitly address DEEP-C, but we have added a comment in our manuscript in relation to Liu et al. [2020] addressing specifically the use of reanalysis in a previous version of the DEEP-C dataset. It is worth noting that almost all our analysis uses only CERES data, and DEEP-C is shown for context.

6. Para 80: SSP2-4.5 is from 2015 onward, not 2014. Please correct.
Done.

7. Para 85: "A positive energy imbalance corresponds to energy leakage, and a negative energy imbalance is associated with an artificial input or source of energy."
Shouldn't it be the opposite? Also, the range is to be expected from natural variability, not something artificial. Please clarify.

You are right that a positive imbalance means a gain of energy over time for a system with constant

forcing that conserves energy. However, climate models do not necessarily conserve energy. For a pre-industrial simulation run to a stationary control state, a positive energy imbalance compensates leakage of energy, and conversely a negative energy imbalance corresponds to a compensation of artificially produced energy inside the model. We have updated the corresponding text to better explain this. Long-term imbalance of -3 to +4 Wm^{-2} for control simulations of 500 years, or longer, is not within the expected range of natural variability on that time scale.

8. Para 115: It would be more appropriate to do S.E. for large ensembles as $\sigma/\text{sqr}t(n)$, so $\text{CI} = \pm 1.64 * \sigma / \text{sqr}t(n)$. Please consider.

It is true that dividing by \sqrt{n} would provide a more robust uncertainty calculation for the models with larger ensembles, in the sense that it would be accurate for the expected forced response of the model. However, in our statistical analysis the use of the S.E would then not represent the full range the model can simulate, and for this reason we use the mean of the standard errors calculated for the individual realizations, in order to ensure meaningful comparisons between models and observations. We have clarified this in the manuscript.

9. Para 120: Units should be $0.1 \text{ Wm}^{-2}\text{decade}^{-1}$, not Wm^{-2} , as it is the observational uncertainty due to drift (temporally varying). Please correct.
Done.

10. Para 125: This is not clear... why not just use the continuous time series? Please clarify. Our reasoning behind the calculations of Section 2.5 is that we aim to investigate the potential effect of the transition between the historical experiment and the future scenario on the overall energy imbalance trend. To this end, we perform linear regression for 20-year periods directly before and after the transition, which shows whether there is a discontinuity in the forcing.

11. Para 160: Fan et al., 2025 too would be a good reference here. Please add. The study by Fan et al. [2025] investigates the impact of using different SST and sea ice datasets, and finds that this can have a surprisingly large impact on the energy imbalance. This problem was also studied by Modak and Mauritsen [2023]. However, the current study is about coupled models, and we found no natural point to include a discussion of atmosphere-only model simulations.

12. Figure 10: Is 12 years enough? especially when 2001-2012 is the heart of the hiatus. Please clarify. We appreciate your comment. In our study we analyze the two cases: short-term and long-term warming. For this first case, we aim to provide a prediction of short-term decadal warming using a 12 year time period. The results are surprisingly promising, but prove to be highly model dependent. Regarding the hiatus period, less observed warming in the first period corresponds to an overall increased temperature change between the two periods, consistent with the magnitude of the energy imbalance in the first period.

13. Fig. 11: What if the y-axis was changed to trend in EEI? How would the results change? Please consider. This was previously investigated by Myhre et al. [2025] showing that there is no consistent relationship between the 2001-2023 energy imbalance trend and climate sensitivity. We have repeated this analysis for the period 2001-2024 and confirmed their results (not shown).

14. Para 260: Since this is the headline result and title perhaps worth quantifying this better? i.e., some correlation or some other statistic rather than just eye-balling Fig. 11. Please consider. After consideration, and after updating our calculations with data for 2025, we have decided to expand the discussion in Section 5 and the conclusions to be more explicit. Due to the inconsistent behaviour of some models due to the identified discontinuity between the historical experiment and

future scenario, an emergent constraint analysis in this case would not provide a reliable constraint for climate sensitivity and as a result would not be so meaningful. For this reason we decided instead to highlight the physical relationship between the magnitude of the energy imbalance and climate sensitivity, utilizing the 2-layer model. We note that the majority of models in fact follow this theoretical relationship, which we believe strengthens our argument.

15. Para 290: Didn't you rule out incorrectly applied forcing? Please clarify.

The paragraph refers to for example incorrectly applied anthropogenic aerosol forcing, which we assume is still one of the potential explanations of the discrepancy between the observations and CMIP6 models. For comparison, the aerosol forcing that we consider in Section 4 is volcanic aerosol forcing. We conclude that the handling of this volcanic forcing, in the transition between the historical experiment and the SSP2-4.5 scenario, cannot explain the overall underestimation of the energy imbalance trend in CMIP6 models for the period 2001-2025.

We have added a clarifying comment in our manuscript.

16. Internal variability alone cannot explain the positive trend in EEI (Raghuraman et al., 2021) so this paragraph needs to be revised with this being restated/acknowledged. Please consider.

Raghuraman et al. [2021] considers purely internal variability as simulated by models in control simulations and finds that this alone cannot explain the trend in EEI. However, when models are realistically forced as in the combined historical and future projection simulations we do find some of the models are consistent as the observed trend is within their simulated internal variability (Figure 6).

References

- Fan, C., Paynter, D., Kramer, R. J., and Lin, P. (2025). Sensitivity of Earth's radiation budget to lower boundary condition data sets in historical climate simulations. *Geophysical Research Letters*, 52(17), e2025GL115914.
- Hansen, J. et al. ,Climate Response Times: Dependence on Climate Sensitivity and Ocean Mixing.Science229,857-859(1985).DOI:10.1126/science.229.4716.857.
- Loeb, N. G., Mayer, M., Kato, S., Fasullo, J. T., Zuo, H., Senan, R., ... and Balmaseda, M. (2022). Evaluating twenty-year trends in Earth's energy flows from observations and reanalyses. *Journal of Geophysical Research: Atmospheres*, 127(12), e2022JD036686.
- Mauritsen, T. et al. (2012), Tuning the climate of a global model, *J. Adv. Model. Earth Syst.*, 4, M00A01, doi:10.1029/2012MS000154.
- Modak, A. and Mauritsen, T.: Better-constrained climate sensitivity when accounting for dataset dependency on pattern effect estimates, *Atmos. Chem. Phys.*, 23, 7535–7549, <https://doi.org/10.5194/acp-23-7535-2023>, 2023.
- Park, C., and Soden, B. J. (2025). Negligible contribution from aerosols to recent trends in Earth's energy imbalance. *Science advances*, 11(48), eadv9429.
- Raghuraman, S. P., Paynter, D., Menzel, R., and Ramaswamy, V. (2023). Forcing, cloud feedbacks, cloud masking, and internal variability in the cloud radiative effect satellite record. *Journal of Climate*, 36(12), 4151-4167.
- von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D., Champollion, N., ... and Wild, M. (2016). An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, 6(2), 138-144.

References

- Timothy Andrews, Jonathan M. Gregory, David Paynter, Levi G. Silvers, Chen Zhou, Thorsten Mauritsen, Mark J. Webb, Kyle C. Armour, Piers M. Forster, and Holly Titchner. Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophysical Research Letters*, 45(16):8490–8499, 2018. doi: <https://doi.org/10.1029/2018GL078887>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078887>.
- Svante Arrhenius. Xxxi. on the influence of carbonic acid in the air upon the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(251):237–276, 1896. doi: 10.1080/14786449608620846. URL <https://doi.org/10.1080/14786449608620846>.
- Chongxing Fan, David Paynter, Ryan J. Kramer, and Pu Lin. Sensitivity of earth’s radiation budget to lower boundary condition data sets in historical climate simulations. *Geophysical Research Letters*, 52(17):e2025GL115914, 2025. doi: <https://doi.org/10.1029/2025GL115914>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2025GL115914>. e2025GL115914 2025GL115914.
- J.B.J. Fourier. *Théorie analytique de la chaleur*. Didot, 1822.
- Jean Baptiste Joseph Fourier. *The Analytical Theory of Heat*. Cambridge University Press, 1878. Reprinted by Cambridge University Press, 2009. Original French publication: 1822.
- J. M. Gregory and T. Andrews. Variation in climate sensitivity and feedback parameters during the historical period. *Geophysical Research Letters*, 43(8):3911–3920, 2016. doi: <https://doi.org/10.1002/2016GL068406>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068406>.
- J. Hansen, G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone. Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, 229:857–859, 1985. doi: 10.1126/science.229.4716.857.
- Chunlei Liu, Richard P. Allan, Michael Mayer, Patrick Hyder, Damien Desbruyères, Lijing Cheng, Jianjun Xu, Feng Xu, and Yu Zhang. Variability in the global energy budget and transports 1985–2017. *Climate Dynamics*, 55(11):3381–3396, 2020. doi: 10.1007/s00382-020-05451-8. URL <https://doi.org/10.1007/s00382-020-05451-8>.
- Thorsten Mauritsen, Bjorn Stevens, Erich Roeckner, Traute Crueger, Monika Esch, Marco Giorgetta, Helmuth Haak, Johann Jungclaus, Daniel Klocke, Daniela Matei, Uwe Mikolajewicz, Dirk Notz, Robert Pincus, Hauke Schmidt, and Lorenzo Tomassini. Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*, 4(3), 2012. doi: <https://doi.org/10.1029/2012MS000154>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012MS000154>.
- A. Modak and T. Mauritsen. Better-constrained climate sensitivity when accounting for dataset dependency on pattern effect estimates. *Atmospheric Chemistry and Physics*, 23(13):7535–7549, 2023. doi: 10.5194/acp-23-7535-2023. URL <https://acp.copernicus.org/articles/23/7535/2023/>.
- Gunnar Myhre, Øivind Hodnebrog, Norman Loeb, and Piers M. Forster. Observed trend in earth energy imbalance may provide a constraint for low climate sensitivity models. *Science*, 388(6752): 1210–1213, 2025. doi: 10.1126/science.adt0647.
- Shiv Priyam Raghuraman, David Paynter, and V. Ramaswamy. Anthropogenic forcing and response yield observed positive trend in earth’s energy imbalance. *Nature Communications*, 12:4577, 07 2021. doi: 10.1038/s41467-021-24544-4.
- Chen Zhou, Mark D. Zelinka, and Stephen A. Klein. Impact of decadal cloud variations on the earth’s energy budget. *Nature Geoscience*, 9(12):871–874, 2016. URL <https://doi.org/10.1038/ngeo2828>.