

Multi-objective observational constraint of tropical Atlantic and Pacific low cloud variability narrows uncertainty in cloud feedback

Corresponding Author: Professor Hui Su

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This paper examined the relationship between tropical marine low cloud feedback with local (the tropical Pacific and Atlantic) and remote (represented by lower-tropospheric temperatures) sea surface temperatures. The authors applied a Pareto optimization approach to observational constraints from both the Pacific and Atlantic, with the latter favoring larger positive feedback. Although I find the analysis very interesting and useful for the current debate in the climate community regarding the levels of climate sensitivity (specifically ECS values) that are seen both in nature and as represented in CMIP models, I believe there are many areas in the paper that need improvement and reconsideration.

I provide some main comments here with more minor comments below:

- My main concern is that the authors take results they found from analysis in two specific subregions of tropical Pacific and Atlantic oceans, and linearly extrapolate their findings and relate them to global ECS values. As their abstract states, "...and suggests a 71% increase in the tropical shortwave cloud feedback ...", this statement is valid based directly on the evidence presented in this paper, however the following bit of text, "...yielding higher climate sensitivity", is based on various assumptions that are not based on evidence. I recommend the authors reconsider the conclusion they are reaching (and therefore reconsider the title of the paper), and perhaps limit the scope of the analysis to the tropical shortwave cloud feedback being investigated instead of extrapolating the findings to suggest higher global ECS values. I provide more comments below on these potentially problematic "assumptions" made in this paper.

L214-232: I indicate here many of the assumptions that I felt are related to the point I made above:

*Line 214: "Assuming the effective radiative forcing and all other feedbacks are independent",

*Line 220: "...hence the best estimate of multi-model ECS should increase by 14%, from 3.76 K in the selected 28 CMIP6 models to 4.32 K."

*Line 223: "An observational constraint favoring more positive tropical SWCF may potentially increase the multi-model extratropical SWCF simultaneously"

*Line 232: "If we assume a physical link between the tropical and extratropical SWCF".

*Line 234: "... leading to an 28% increase in the best estimate of multi-model ECS to 4.81 K."

You are assuming linearity in this whole approach. And you are assuming that your estimates of how local SSTs and T700 impact cloud feedback is representative of the whole globe. I think you need to scale back these assumptions and conclusions. You are ignoring all other processes that can have an impact on global ECS. For example, this analysis considers two select sub-regions in the tropics, not the whole tropics. Furthermore, there are many regions of the globe (e.g. cryosphere? forests?) that can impact global climate sensitivity and are not even considered. For example, vegetation influences ECS through various processes, such as altering the surface albedo (reflectivity), affecting the water cycle, and influencing the exchange of greenhouse gases like carbon dioxide and methane between the atmosphere and the biosphere. The cryosphere can impact ECS through albedo, ocean circulation, greenhouse gas release, and sea level rise. These are just a few processes that can impact global ECS, and the fact that the authors ignore all of these and make a linear extrapolation from the analysis done here to global ECS values is, I think, incorrect. The authors can at least discuss how these assumptions can impact their results and conclusions.

- My second main comment is related to how difficult I found reading many parts of the paper. This I think is mainly due to the extensive use of acronyms throughout the paper. Although I understand the nature of writing manuscripts like this requires the use of acronyms in many places, I found it very difficult to smoothly read through the paper without having to scroll back and forth to remember what various acronyms meant. To help with this, perhaps the authors can consider including a table that defines the acronyms? It is possible this might not be a problem for others, but for me it was very difficult to read, so I recommend the authors take this into account for the sake of the community of readers.

Minor comments:

Line 55: These observational constraints are not described in detail. Where is this information coming from? The information might be in the methods section, but I suggest that the authors provide at least 1-2 sentences about the observations here for better context.

Line 64: Of the two largest? The two largest?

Line 88: Here and other places that mention the error bars: How is this standard error defined? Is this explained in the methods section? More details are needed.

Line 100: Not sure this is true. The subensembles are by definition bounded by the envelope of the original models, unless you allow for combinations outside the convex hull of the model space, which I don't think is true in this case. Therefore, it is not accurate to make the conclusion of the linearity of the SWCRE sensitivities based on this fact.

L197: Do you have enough data points in Figure 4b to create and describe this distribution thoroughly (28 models total)? Figure 4a addresses this point well, and I am not sure why Figure 4b is even included? I recommend removing Figure 4b from the text or moving it to the supplement unless there is a clear reason for it.

-Line 223: Here is another assumption that is not validated with evidence. Figure 6 shows evidence that some models have a tropical-extratropical relationship that is contrary to this statement (i.e. models in the 4th quadrant of figure 6).

-Line 256: Why do the subensembles only include 5 models per ensemble and not all possible model combinations, as was done in Massoud et al. 2023? That paper used Bayesian Model Averaging to constrain ECS in CMIP6 models and found the opposite of the conclusions on this study to be true (lower overall ECS was more likely and resulted in future temperature projections that matched other studies). I recommend the authors look at that paper and consider incorporating the findings from there into this discussion for comparison. Another work that is already cited in this paper is Hausfather et al. 2022 which makes similar conclusions to Massoud et al., 2023. The authors grab the ECS values from Hausfather et al., 2022, and do not consider or mention that the conclusions made in that paper are opposite to their own. I highly recommend the authors consider the global picture when making assumptions about global ECS.

Paper: Massoud et al., 2023: <https://doi.org/10.1038/s43247-023-01009-8>

Line 273: "... CMIP6 models ... systematically underestimate the cloud sensitivities to local SST." Perhaps this should be the main conclusion of this paper, instead of the assumption that this is linearly extrapolated to indicate higher global ECS values.

Figure 3: Panel b is not described in the caption. Unless I missed something.

Line 356: ERA5 is not defined.

Reviewer #2

(Remarks to the Author)

In this study, the authors use a cloud-controlling factor analysis based on SST and the temperature at 700 hPa to constrain global SW low cloud feedback in CMIP6 climate models using CERES and MODIS observations. For that purpose, they further sub sample the CMIP6 model ensemble into 5-model ensembles and use a statistical approach to constrain the sensitivity of the simulated low clouds to these two controlling factors in two stratocumulus regions. They found that the sub-ensembles that agree best with observations correspond to a higher climate sensitivity likelihood than the multimodel mean.

I find this study interesting and well written for the most part. I have some concerns about the approach that I further develop below. Furthermore, while I think the study is complementing the existing literature, it seems to be more of an incremental improvement rather than a substantial advance of the field. We already know that the low cloud feedback is very likely positive and there is a pretty good constraint on what it should be and how it would affect ECS. These papers are cited by the authors. Then the main novelty relies on the fact that the authors' results suggest slightly higher multimodel mean ECS than the other studies, yet, this might be biased by the model sub-sampling. Here are my main comments:

T700 argument.

I'm not convinced by using T700 instead of EIS (or ECTEI) since EIS captures the effect of warming coming from tropical ascent as largely discussed by Qu et al. (2015) and is directly linked to low cloud formation through physical mechanisms (e.g., Wood and Bretherton, 2006). It's not even clear if T700 is well correlated with low-cloud fraction (the correlation

between T700 ascent and local is not that clear in the sc regions as well in Fig. 1c). Somewhat related to this, the argument about ECTEI not working well in some region could be simply explained by the fact that controlling factors have different influences depending on the region and the omission of one controlling factor that is important in that region could very well explained the inconsistencies between the two regions. For example, the wind speed is relatively important in some seasons (e.g., Naud et al., 2023)

Statistical approach.

I'm not a specialist of the approach that is used and so I won't comment on it, however, 28 members and 5 sub-ensembles are used, which should correspond to 98280 combinations if I'm not mistaken. How do you get to ~ 200K? Also, nothing is said about the sensitivity to the sub-ensemble number you pick. Why choosing 5?

Extratropical and Tropical cloud feedback link.

I'm not sure what type of link must exist between tropical and extratropical feedback to allow the use of the Pareto constraint, but the link is clearly not linear from Fig. 6, which makes me wonder how valid the subsequent constraint is.

Influence of CESM-based models.

Then another main concern is the substantial impact of CESM-based models on the sampling. Six of the seven high-ECS models on the upper right corner in Fig. 6 are CESM-based models, among which four originate from the same CESM2 model and another one is very similar to CESM2 (E3SM). What would happen there was only one CESM2 model? How would this affect the sub-ensembles and the whole analysis?

Cloud feedback computations.

The feedback computation is not conventional. The authors are using the first and last 30 years of the 4xCO₂ experiment instead of comparing the 4xCO₂ to piControl experiments. I can understand the logic of doing that but there are rapid adjustments of the clouds in the first few years of the 4xCO₂ simulations (Zelinka et al., 2013). This makes it difficult to compare with existing literature and it also weakens the validity of their computations using existing feedback from previous literature in the section "Implications for ECS". It would be more convincing to see some comparison between classic feedbacks and the ones used by the authors, and also to see whether removing the first few (~3) years would affect these 4xCO₂-based feedback results.

References:

Naud, C. M., Elsaesser, G. S., & Booth, J. F. (2023). Dominant Cloud Controlling Factors for Low-Level Cloud Fraction: Subtropical Versus Extratropical Oceans. *Geophysical Research Letters*, 50(19). <https://doi.org/10.1029/2023GL104496>

Qu, X., Hall, A., Klein, S. A., & Caldwell, P. M. (2015). The strength of the tropical inversion and its response to climate change in 18 CMIP5 models. *Climate Dynamics*, 45(1–2), 375–396. <https://doi.org/10.1007/s00382-014-2441-9>

Zelinka, M. D., Klein, S. A., Taylor, K. E., Andrews, T., Webb, M. J., Gregory, J. M., & Forster, P. M. (2013). Contributions of different cloud types to feedbacks and rapid adjustments in CMIP5. *Journal of Climate*, 26(14), 5007–5027. <https://doi.org/10.1175/JCLI-D-12-00555.1>

Reviewer #3

(Remarks to the Author)

This paper created a so-called emergent constraint on the climate sensitivity using low cloud feedbacks over the tropical Pacific and Atlantic, and the results suggest a higher climate sensitivity.

This new emergent constraint is very similar to that created by Myers et al. 2021, but the results of these two studies are distinct. The authors of this paper suggest that the different is induced by choices in cloud controlling factors, but the linear combination of SST and T700 partial derivatives in this study is actually equivalent to the combination of SST and LTS, due to the linear correlation between SST, T700 and LTS. In Myers et al. 2021, EIS is chosen instead of LTS. It is well known that EIS is better predictor than LTS in the tropics, so the choice of cloud controlling factors in this paper is not as good as that in Myers et al. 2021. Therefore, the method used in this paper might not be novel enough for Nature Communications.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed all of my previous comments, and I think the paper is much improved. The inclusion of Figure 4 provides a clear spatial representation of the regional constraints and helps give context to the overall message of the paper. The emphasis on the CMIP models' challenges in simultaneously capturing these trade-off dynamics is a valuable contribution to the field and will benefit the broader community. I recommend the paper for publication.

Reviewer #2

(Remarks to the Author)

This is the second round of review for Wu et al.

I appreciate the efforts made by the authors to address my concerns. IMO, the manuscript has been improved. There are still two points I'd like the authors to consider:

1) Regarding the choice of T700 vs. EIS. It is true that the T700 is a more direct measure of remote influence when it comes to a simple linear regression between cloud amount and T700 vs. EIS, however, for bilinear regressions, since SST is held constant, then most of the EIS variability (since RH is often assumed to be 80% in EIS computations) then comes from T700 variability... As a result, one would expect similar results when looking at a 2D diagram using $\partial C/\partial EIS$ as a function $\partial C/\partial SST$ instead of $\partial C/\partial T700$. So I would recommend making this clear in the text in the discussion section at least, when comparing against previous literature.

2) Regarding the feedbacks computations, it is not clearly explained. What radiative kernels did the author use and, more importantly, did they compute these radiative kernel feedbacks using the first and last 30 years of the 4xCO₂ experiments? My point was to compare traditional feedbacks, which are computed by comparing 4xCO₂ with piControl runs, rather than the first and last 30 years of 4xCO₂ runs. Glad to see that radiative kernels vs $\Delta SWCRE/\Delta T$ are similar though.

Reviewer #3

(Remarks to the Author)

The quality of the draft has been improved after the authors revised the paper, but the novelty of this paper remains unchanged. As reviewer #2 pointed out, the study is more of an incremental improvement rather than a substantial advance of the field. Therefore, I think this paper is more suitable for Journal of Climate.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Overview of response to Reviewers:

First, we want to thank the reviewers for their constructive comments. In response to the reviewer's comments, we have removed to better highlight the novel aspects of this study: observational constraints for tropical marine stratocumulus feedback that reconcile difficulties that models have in simultaneously simulating the Atlantic and Pacific regions. We have updated our title to reflect this. The reviewers' comments motivated the addition of a new Fig. 4 to provide a spatial visualization of the regional constraints, while helping to illustrate the importance of reconciling these constraints. The linear inference of ECS from tropical shortwave cloud feedback has been removed because, while it aimed to put the tropical low-cloud feedback results in a larger context, it created a distraction from the main points. We have also further clarified the reason for choosing T_{700} to represent the marine stratocumulus response to ascent area warming with Supplementary Figure 7.

The point-by-point response to reviewer's comments is given below.

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

This paper examined the relationship between tropical marine low cloud feedback with local (the tropical Pacific and Atlantic) and remote (represented by lower-tropospheric temperatures) sea surface temperatures. The authors applied a Pareto optimization approach to observational constraints from both the Pacific and Atlantic, with the latter favoring larger positive feedback. Although I find the analysis very interesting and useful for the current debate in the climate community regarding the levels of climate sensitivity (specifically ECS values) that are seen both in nature and as represented in CMIP models, I believe there are many areas in the paper that need improvement and reconsideration.

I provide some main comments here with more minor comments below:

- My main concern is that the authors take results they found from analysis in two specific sub regions of tropical Pacific and Atlantic oceans, and linearly extrapolate their findings and relate them to global ECS values. As their abstract states, "...and suggests a 71% increase in the tropical shortwave cloud feedback ...", this statement is valid based directly on the evidence presented in this paper, however the following bit of text, "...yielding higher climate sensitivity", is based on various assumptions that are not based on evidence. I recommend the authors reconsider the conclusion they are reaching (and therefore reconsider the title of the paper), and perhaps limit the scope of the analysis to the tropical shortwave cloud feedback being investigated instead of extrapolating the findings to suggest higher global ECS values. I provide more comments below on these potentially problematic "assumptions" made in this paper.

L214-232: I indicate here many of the assumptions that I felt are related to the point I made above:

*Line214: "Assuming the effective radiative forcing and all other feedbacks are independent",

*Line 220: "...hence the best estimate of multi-model ECS should increase by 14%, from 3.76 K in the selected 28 CMIP6 models to 4.32 K."

*Line 223: "An observational constraint favoring more positive tropical SWCF may potentially increase the multi-model extratropical SWCF simultaneously"

*Line 232: "If we assume a physical link between the tropical and extratropical SWCF".

*Line 234: "... leading to a 28% increase in the best estimate of multi-model ECS to 4.81 K."

We agree with the reviewer that our inference about ECS relies on several unsubstantiated assumptions. Following the reviewer's suggestion, now we emphasize more on the tropical shortwave cloud feedback in the title, the abstract and the main text, and have removed the part of a quantitative ECS inference.

You are assuming linearity in this whole approach. And you are assuming that your estimates of how local SSTs and T700 impact cloud feedback is representative of the whole globe. I think you need to scale back these assumptions and conclusions. You are ignoring all other processes that can have an impact on global ECS. For example, this analysis considers two select sub-regions in the tropics, not the whole tropics. Furthermore, there are many regions of the globe (e.g. cryosphere? forests?) that can impact global climate sensitivity and are not even considered. For example, vegetation influences ECS through various processes, such as altering the surface albedo (reflectivity), affecting the water cycle, and influencing the exchange of greenhouse gases like carbon dioxide and methane between the atmosphere and the biosphere. The cryosphere can impact ECS through albedo, ocean circulation, greenhouse gas release, and sea level rise. These are just a few processes that can impact global ECS, and the fact that the authors ignore all of these and make a linear extrapolation from the analysis done here to global ECS values is, I think, incorrect. The authors can at least discuss how these assumptions can impact their results and conclusions.

In addition to the abovementioned changes, we also add a new paragraph discussing the other important processes that can influence ECS, such as vegetation and cryosphere mentioned by the reviewer (Lines 224-228).

- My second main comment is related to how difficult I found reading many parts of the paper. This I think is mainly due to the extensive use of acronyms throughout the paper. Although I understand the nature of writing manuscripts like this requires the use of acronyms in many places, I found it very difficult to smoothly read through the paper without having to scroll back and forth to remember what various acronyms meant. To help with this, perhaps the authors can consider including a table that defines the acronyms? It is possible this might not be a problem for others, but for me it was very difficult to read, so I recommend the authors take this into account for the sake of the community of readers.

We have added an acronym table.

Minor comments:

Line 55: These observational constraints are not described in detail. Where is this information coming from? The information might be in the methods section, but I suggest that the authors provide at least 1-2 sentences about the observations here for better context.

In the previous version, the description of the observational data was given in the "Data and Methods" section (Lines 348-360). Now we add one sentence in the introduction to briefly summarize the observational data used in this study (Lines 49-51).

Line 64: Of the two largest? The two largest?

We have corrected the phrase as "the two largest".

Line 88: Here and other places that mention the error bars: How is this standard error defined? Is this explained in the methods section? More details are needed.

Now we have clarified in the main text that the standard error is computed from the ordinary-least-square linear regression (Lines 90-92).

Line 100: Not sure this is true. The subensembles are by definition bounded by the envelope of the original models, unless you allow for combinations outside the convex hull of the model space, which I don't think is true in this case. Therefore, it is not accurate to make the conclusion of the linearity of the SWCRE sensitivities based on this fact.

We have removed this sentence, which does not influence our main analysis and argument. Instead, we add a new Supplementary Table 1 containing results of t-tests and F-tests to justify our linear regressions.

L197: Do you have enough data points in Figure 4b to create and describe this distribution thoroughly (28 models total)? Figure 4a addresses this point well, and I am not sure why Figure 4b is even included? I recommend removing Figure 4b from the text or moving it to the supplement unless there is a clear reason for it.

We agree with the reviewer that the sample size is too small in Fig. 4b, and we have removed it.

-Line 223: Here is another assumption that is not validated with evidence. Figure 6 shows evidence that some models have a tropical-extratropical relationship that is contrary to this statement (i.e. models in the 4th quadrant of figure 6).

We have removed Fig. 6, and just cite relevant papers to argue a possible tropical-extratropical connection which was already reported in the literature (Lines 220-223).

-Line 256: Why do the subensembles only include 5 models per ensemble and not all possible model combinations, as was done in Massoud et al. 2023? That paper used Bayesian Model Averaging to constrain ECS in CMIP6 models and found the opposite of the conclusions on this study to be true (lower overall ECS was more likely and resulted in future temperature projections that matched other studies). I recommend the authors look at that paper and consider incorporating the findings from there into this discussion for comparison. Another work that is already cited in this paper is Hausfather et al. 2022 which makes similar conclusions to Massoud et al., 2023. The authors grab the ECS values from Hausfather et al., 2022, and do not consider or mention that the conclusions made in that paper are opposite to their own. I highly recommend the authors consider the global picture when making assumptions about global ECS.

Paper: Massoud et al., 2023: <https://doi.org/10.1038/s43247-023-01009-8>

We have added a new Supplementary Figure 2 and one sentence in the main text to justify the subensemble size of 5 (Lines 99-101).

We have also cited the two papers recommended by the reviewer when discussing processes other than marine stratocumulus feedback that may cause a positive ECS bias in CMIP6 models.

Line 273: "... CMIP6 models ... systematically underestimate the cloud sensitivities to local SST." Perhaps this should be the main conclusion of this paper, instead of the assumption that this is linearly extrapolated to indicate higher global ECS values.

We have removed the linear extrapolation estimate of ECS and accepted the reviewer's suggestion on the main conclusion.

Figure 3: Panel b is not described in the caption. Unless I missed something.

The missing caption is now put back.

Line 356: ERA5 is not defined.

We have added a full name of ERA5 (Lines 357-358).

Reviewer #2 (Remarks to the Author):

In this study, the authors use a cloud-controlling factor analysis based on SST and the temperature at 700 hPa to constrain global SW low cloud feedback in CMIP6 climate models using CERES and MODIS observations. For that purpose, they further sub sample the CMIP6 model ensemble into 5-model ensembles and use a statistical approach to constrain the sensitivity of the simulated low clouds to these two controlling factors in two stratocumulus regions. They found that the sub-ensembles that agree best with observations correspond to a higher climate sensitivity likelihood than the multimodel mean.

I find this study interesting and well written for the most part. I have some concerns about the approach that I further develop below. Furthermore, while I think the study is complementing the existing literature, it seems to be more of an incremental improvement rather than a substantial advance of the field. We already know that the low cloud feedback is very likely positive and there is a pretty good constraint on what it should be and how it would affect ECS. These papers are cited by the authors. Then the main novelty relies on the fact that the authors' results suggest slightly higher multimodel mean ECS than the other studies, yet, this might be biased by the model sub-sampling. Here are my main comments:

In the revised version, we further highlighted our multi-objective optimization approach as the novelty of the paper. When considering the different low cloud sensitivities to local and remote SST warming in the Pacific and Atlantic stratocumulus regions, none of the CMIP6 models performs well in both regions simultaneously. Our multi-objective optimization approach allows for an objective weighting between the two regions to handle the trade-off. The new Fig. 4 illustrates how the Pareto optimization approach reduces both the error and the uncertainty range relative to the original CMIP6 model ensemble. Compared with traditional emergent constraints, our observational constraint derived from this approach does not depend on hunting for a linear relationship across the model ensemble. We also further clarify the advantage of choosing SST and T_{700} as the two CCFs representing the effects of local and remote SST warming, which is explained in detail in the following responses.

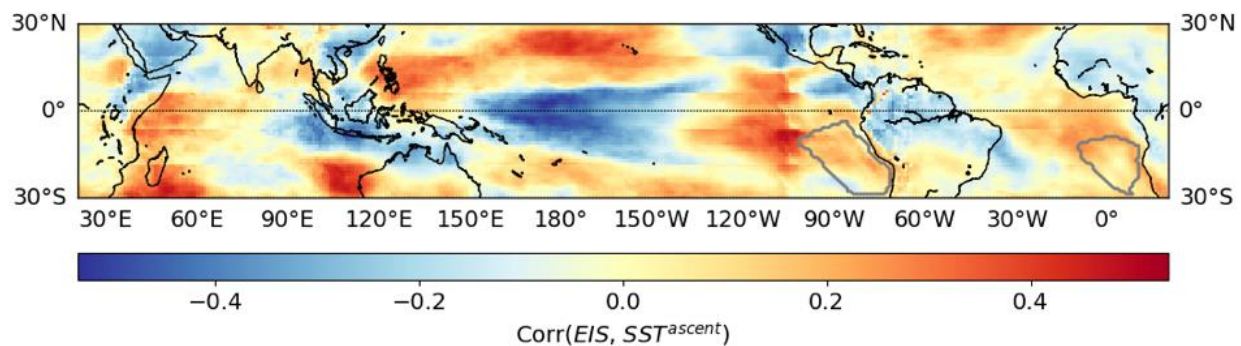
T700 argument.

I'm not convinced by using T700 instead of EIS (or ECTEI) since EIS captures the effect of warming coming from tropical ascent as largely discussed by Qu et al. (2015) and is directly linked to low cloud formation through physical mechanisms (e.g., Wood and Bretherton, 2006). It's not even clear if T700 is well correlated with low-cloud fraction (the correlation between T700ascent and local is not that clear in the sc regions as well in Fig. 1c).

We thank the reviewer for this comment as this point has helped us to revise to better communicate our arguments. EIS has an approximate physical interpretation in terms of local boundary layer processes, but

its variation can be caused by both local SST and T_{700} (remote warming). Rather than using a single metric like EIS or ECTEI that condenses information, we examine more CCFs in a multivariate linear regression (Fig. 1) to diagnose physical pathways of marine low cloud formation, including the remote pathway that is relevant to the SST pattern effect.

In this sense, a single CCF does not have to show a high univariate correlation with low clouds. In particular, T_{700} is the agent of the remote pathway, but it can only influence low clouds together with local SST. The decomposition from Fig. 1 suggests that T_{700} dominates the influence of remote SST in ascent regions over low clouds among all the CCFs, which is consistent with a tropical weak temperature gradient and the cloud effects of EIS. By contrast, EIS involves a competition between SST and T_{700} and thus blends the effects of local and remote SST warming. To support our argument, we also produce the EIS version of Fig. 1 as Supplementary Figure 7.



The attached panel from Supplementary Figure 7 shows that EIS does not show a consistent positive correlation with remote tropical SST, making the remote pathway less clear than T_{700} .

Somewhat related to this, the argument about ECTEI not working well in some region could be simply explained by the fact that controlling factors have different influences depending on the region and the omission of one controlling factor that is important in that region could very well explain the inconsistencies between the two regions. For example, the wind speed is relatively important in some seasons (e.g., Naud et al., 2023)

We have added discussion and citations as the reviewer suggests. A stronger SST sensitivity in the Atlantic is likely to be associated with a larger surface latent heat flux term (Qu et al., 2014) and the wind speed plays a less important role than the surface relative humidity (Supplementary Text 1).

Statistical approach.

I'm not a specialist of the approach that is used and so I won't comment on it, however, 28 members and 5 sub-ensembles are used, which should correspond to 98280 combinations if I'm not mistaken. How do you get to ~200K? Also, nothing is said about the sensitivity to the sub-ensemble number you pick. Why choosing 5?

In our 5-member subensemble, any model is allowed to repeat for any times, and therefore the number of combinations will be much larger. If no repetition is allowed, then 98280 will be the correct result. A limited subensemble size is due to the computational cost of the Pareto optimization. A subensemble size of 5 has been shown sufficient while computationally reasonable for Pareto optimization in Langenbrunner & Neelin (2017) and in Supplementary Figure 2.

Extratropical and Tropical cloud feedback link.

I'm not sure what type of link must exist between tropical and extratropical feedback to allow the use of the Pareto constraint, but the link is clearly not linear from Fig. 6, which makes me wonder how valid the subsequent constraint is.

In the revised version, now we only briefly discuss possible implications of our results on extratropical cloud feedback and replaced Fig. 6 by relevant citations from literature (Lines 220-223).

Influence of CESM-based models.

Then another main concern is the substantial impact of CESM-based models on the sampling. Six of the seven high-ECS models on the upper right corner in Fig. 6 are CESM-based models, among which four originate from the same CESM2 model and another one is very similar to CESM2 (E3SM). What would happen there was only one CESM2 model? How would this affect the sub-ensembles and the whole analysis?

We have reproduced the Pareto optimization analysis after excluding all CESM2 models as Supplementary Figures 5 and 6. Although the whole data cloud looks a bit different in the figure, the conclusions based on the Pareto optimal set are insensitive to this exclusion.

Cloud feedback computations.

The feedback computation is not conventional. The authors are using the first and last 30 years of the 4xCO₂ experiment instead of comparing the 4xCO₂ to piControl experiments. I can understand the logic of doing that but there are rapid adjustments of the clouds in the first few years of the 4xCO₂ simulations (Zelinka et al., 2013). This makes it difficult to compare with existing literature and it also weakens the validity of their computations using existing feedback from previous literature in the section "Implications for ECS". It would be more convincing to see some comparison between classic feedbacks and the ones used by the authors, and also to see whether removing the first few (~3) years would affect these 4xCO₂-based feedback results.

We add Supplementary Figure 4 displaying a very good agreement between our simplified feedback computation and more accurate results from a radiative kernel (Zelinka et al., 2020).

Reviewer #3 (Remarks to the Author):

This paper created a so-called emergent constraint on the climate sensitivity using low cloud feedbacks over the tropical Pacific and Atlantic, and the results suggest a higher climate sensitivity.

This new emergent constraint is very similar to that created by Myers et al. 2021, but the results of these two studies are distinct. The authors of this paper suggest that the difference is induced by choices in cloud controlling factors, but the linear combination of SST and T700 partial derivatives in this study is actually equivalent to the combination of SST and LTS, due to the linear correlation between SST, T700 and LTS. In Myers et al. 2021, EIS is chosen instead of LTS. It is well known that EIS is better predictor than LTS in the tropics, so the choice of cloud controlling factors in this paper is not as good as that in Myers et al. 2021. Therefore, the method used in this paper might not be novel enough for Nature Communications.

We thank the reviewer for pointing out that a hurried reader might not immediately spot the difference between the approach taken here and previous approaches that use observations and regressions on cloud controlling factors. We indeed greatly appreciate the Myers et al. 2021 paper (M21, duly cited), but the approach is really quite different and we have revised to clarify this.

We outline in the following paragraphs the specific revisions we have made for this. But first to recap at a higher-level: Overall the revisions aim to better highlight the novel aspects of this study: observational constraints for tropical marine stratocumulus feedbacks that reconcile difficulties that models have in simultaneously simulating the Atlantic and Pacific regions. We hope that the addition of a new Fig. 4 with a spatial visualization of the regional constraints better illustrates the importance of reconciling such constraints.

M21 does not evaluate the climate models explicitly for the current climate feedbacks, which is key to our approach. As a result, they could not see the difficulties that the models have in correctly capturing the observed shortwave response to cloud controlling factors in major stratocumulus regions — Atlantic and Pacific — simultaneously. Identifying and reconciling this difficulty is our major emphasis — which motivated us to apply the Pareto optimization to a multi-objective observational constraint.

We have revised to bring this out starting from the abstract (Lines 14-17) and have moved the first mention of T_{700} and SST to the results section so it does not distract early on. More importantly, we have included a new Figure 4 that makes the spatial aspects that we are addressing in the observations come out more clearly. Relative to simple multi-model means, our Pareto-optimal mean of the CMIP6 models can both reduce the model error and its standard deviation with a clear spatial pattern. We have also highlighted the advantage of our combination of Pareto optimization and Bayesian a priori observational constraints, which is a major part of the novelty of our paper.

For the reviewer's interest, we also note that M21 effectively assumes that the sensitivity to cloud controlling factors is the same in current climate observations as in the models on global warming timescales (hence avoiding the need to evaluate the model's interannual variability). Although the apparent difference in conclusion is actually from the role of cumulus clouds in their global analysis, it is encouraging that results regarding the CMIP6 stratocumulus feedback are consistent between the approaches in that CMIP6 models tend to underestimate it compared with observations.

Regarding the choice of cloud controlling factors, we of course began by examining EIS and ECTEI which condense information from multiple meteorological variables. It would have been lovely if we could simply have used one of these. However, because of the trade-offs between the Atlantic and Pacific regions, which are not equally well captured by these traditional compacting measures (Fig. 2), we redid the cloud controlling factor analysis with multiple factors. This evaluation led to the choice of T_{700} and SST, which also give more physical insight into the local and remote pathways (Fig. 1). As is addressed in our responses to the second reviewer, T_{700} or LTS do not have to be a better CCF than EIS because our multivariate regression has included all the necessary CCFs. However, T_{700} better represents the influence of remote SST warming in ascent regions, which is directly related with the SST pattern effect. For comparison, we have included a new Supplementary Figure 7 that shows the EIS-SST evaluation explicitly. Unlike T_{700} , EIS does not show a consistent positive correlation with tropical remote SST, and the remote pathway cannot be fully represented without T_{700} . We have now clarified the wider examination of cloud controlling factors, beginning in the introduction Lines 49-50 and continuing in results Lines 70-84 and Fig. 1.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed all of my previous comments, and I think the paper is much improved. The inclusion of Figure 4 provides a clear spatial representation of the regional constraints and helps give context to the overall message of the paper. The emphasis on the CMIP models' challenges in simultaneously capturing these trade-off dynamics is a valuable contribution to the field and will benefit the broader community. I recommend the paper for publication.

Reviewer #1 (Remarks on code availability):

By providing a notebook-style environment to run the code with detailed explanation, the authors have done a sufficient job in providing the code to run this analysis.

We thank the reviewer for their constructive comments in the previous review, which helped us improve the manuscript.

Reviewer #2 (Remarks to the Author):

This is the second round of review for Wu et al.

I appreciate the efforts made by the authors to address my concerns. IMO, the manuscript has been improved. There are still two points I'd like the authors to consider:

1) Regarding the choice of T700 vs. EIS. It is true that the T700 is a more direct measure of remote influence when it comes to a simple linear regression between cloud amount and T700 vs. EIS, however, for bilinear regressions, since SST is held constant, then most of the EIS variability (since RH is often assumed to be 80% in EIS computations) then comes from T700 variability... As a result, one would expect similar results when looking at a 2D diagram using $\partial C / \partial EIS$ as a function $\partial C / \partial SST$ instead of $\partial C / \partial T700$. So I would recommend making this clear in the text in the discussion section at least, when comparing against previous literature.

We agree with the reviewer that the reconstructed time series of the cloud variable (e.g., SWCRE) are similar from the bivariate linear regression models of SST + EIS or SST + T₇₀₀, because EIS can be approximated as a linear combination of SST and T₇₀₀ when RH is assumed constant. We have added sentences in the discussion section (Lines 244-247) to clarify this together with the new Supplementary Figure 8. However, the regression coefficients are still different in these two linear models and the Pareto optimization directly uses the regression coefficients.

2) Regarding the feedbacks computations, it is not clearly explained. What radiative kernels did the author used and, more importantly, did they compute these radiative kernel feedbacks using the first and last 30 years of the 4xCO₂ experiments? My point was to compare traditional

feedbacks, which are computed by comparing 4xCO₂ with piControl runs, rather than the first and last 30 years of 4xCO₂ runs. Glad to see that radiative kernels vs $\Delta\text{SWCRE}/\Delta T$ are similar though.

We have edited the sentence (Lines 297-300) to clarify that the radiative kernel results were directly extracted from the reference (Zelinka et al., 2020). Zelinka and others indeed compute the radiative feedbacks using the difference between abrupt4xCO₂ and piControl runs. In our Supplementary Figure 4, we attempt to show that a simple estimate using the first and last 30 years of abrupt4xCO₂ runs agrees reasonably well with Zelinka et al.'s radiative kernel results.

Reviewer #2 (Remarks on code availability):

Unfortunately, I do not code in python so I couldn't review the code.

Reviewer #3 (Remarks to the Author):

The quality of the draft has been improved after the authors revised the paper, but the novelty of this paper remains unchanged. As reviewer #2 pointed out, the study is more of an incremental improvement rather than a substantial advance of the field. Therefore, I think this paper is more suitable for Journal of Climate.

We believe our work provides a novel observational constraint for tropical marine stratocumulus feedbacks that objectively reconciles the CMIP models' challenge in simultaneously simulating the Atlantic and Pacific stratocumulus clouds. This constraint is achieved by combining Pareto optimization and Bayesian methods. We are glad to hear positive comments from Reviewers 1 and 2 that our emphasis on this model challenge "is a valuable contribution to the field and will benefit the broader community."

Reviewer #3 (Remarks on code availability):

I didn't find a README file, so I didn't reproduce the dataset of this paper. I believe that all results of this paper are reproducible using the code, because the reputation of the corresponding author is good.

We have created a README file in the Code Ocean capsule to describe how to run our code.