Yin et al (2022): "Climate Impacts of Parameterizing Subgrid Variation and Partitioning of Land Surface Heat Fluxes to the Atmosphere with the NCAR CESM1.2"

Overview

Yin et al. (2022) builds on work presented in *Sun et al.* (2021), which introduced a new parameterization to capture the impacts of subgrid surface fluxes by randomly sampling from distributions of sensible and latent heat fluxes within a gridcell and using those samples to drive unique realizations of the PBL and convection schemes in CESM1.2. *Yin et al.* expand on that to include a condition that the sensible and latent heat fluxes being sampled are either positively or negatively correlated with one another to better capture their relationship and land-atmosphere interactions. The method is novel and holds promise for inclusion in GCMS, though I believe major revisions are needed before this work is published. My main concern is that while model improvements stemming from EXP_COR are strongly emphasized, regions where performance has been degraded are rarely discussed. It is not yet clear that this version of the parameterization represents a strong improvement over the original EXP run. The discussion around each of the figures requires major revision as a result. It would also be beneficial to highlight the model performance globally with more detail.

Major comments

- Section 2.1: Since the parameterization introduced in *Sun et al.* 2021 is still so new and not yet widely known/implemented, further explanation and clarification should be included here. More details of the formulation of the normal distribution (i.e., their Eq. 3) would be helpful at minimum.
- Line 126: The PBL and convective parameterizations are called 16 times per time step to fully sample these subgrid fluxes; what's the impact of that on model efficiency (run time vs. throughput)? Would this be cost prohibitive to implement in the current version of CAM (CAM5) for example, due to its use of CLUBB? It's suggested later that this parameterization could be easily added to CESM and other GCMS, but this information is key to that statement.
- Lines 148-150: I appreciate the need to limit the scope of this study but given that a portion of the motivation for this work was a more through exploration of other variables (2m temperature, surface fluxes, clouds, etc.) and that this is a modified parameterization relative to *Sun et al.* – it would be useful to first confirm that this is still a critical region/season where the new scheme produces a sizable impact.
    - It's noted that global mean values are assessed in Section 3.3, but there are no spatial maps either annually or seasonally, which could strengthen the study.
- Lines 154-155: "It is encouraging that after taking the subgrid energy partitioning into account, the longstanding biases are efficiently mitigated in the EXP_COR run." This is not readily apparent in the bias plots (Fig 2f vs. 2d). Though I see that relative to the CTL

(Fig 2e), there is indeed a small decrease in rain rate along the southern edge of the TP, the bias is still over 5 mm/day. It also appears that a dry bias has been re-introduced over Eastern China in EXP_COR that had been mitigated in EXP. A more thorough discussion of where EXP_COR does *not* improve precipitation is warranted.

- Figure 3: My understanding is that the sum of panels 3a and 3c should match Fig. 2c, and that 3b+3d should equal 2e – is this correct? If so, I wonder if there is a slight plotting mismatch; in Fig 2c, changes in rainfall rarely extend past the TP border, and the small region that does is on the order of 0.5-1 mm/day. But in Fig 3a and 3c, there are precipitation increases that extend along much of the southeastern TP border and into the TP itself, which can be on the order of 1-2 mm/day.
- Lines 200-201: "In contrast, in the EXP_COR run (Fig. 5d), the SLP simulation is corrected…" This does not appear to be the case. The decrease in SLP on the TP in the EXP_COR vs. the CTL case is less than 1 hPa, but the bias in the CTL relative to MERRA2 is more than 10 hPa; the shift is quite small and does not correct the existing bias (though it does reduce it slightly).
- Lines 217-218: "…the underestimations on the southern and eastern margins of the TP in the CTL runs are remarkably improved (Fig. 6f)." When looking for model improvements, it would make more sense to focus on the comparison against observations rather than the control, which would be Fig. 6d. In that light, there is again certainly some improvement in the model bias, but the scale of that change is perhaps minor: on the order of what looks like 2-4 W/m2 when the bias is more than 20 W/m2. The language used in the text suggests a stronger change than is visible.
- Lines 220-221: "The sensible heat flux changes in the EXP_COR run are more significant than those in the EXP run (Fig. 7f), especially over northern China and the southern and eastern margins of the TP, resulting in better agreement with observations." The change over northern China seems to be stronger in the EXP than the EXP_COR simulation, while the increase in SHFLX over eastern China seems to have actually *increased* the bias there relative to both the CTL and EXP cases. Focusing in on the southern edge of the TP may miss out on the bigger regional picture in this case.
- Lines 251-252: "In the EXP_COR run, in addition to the already existing significant improvements in the EXP run…" While it's stated that the EXP run alleviates the overestimation of surface shortwave flux in southern China, it is important to note that the EXP_COR run reduces that improvement and even worsens the bias over eastern China (Fig. 2d). It also seems that the biases in EXP_COR are larger than in EXP over north eastern China/southern Russia, despite general improvements over the TP.
- Figure 12: Unfortunately this diagram is incredibly hard to read. The legend and key need to use a much larger font size to be readable, and even within the plot itself, distinguishing shapes/numbers from each other is challenging. A table may be easier to read, but at minimum the Taylor diagram needs to be heavily revised.
    - Section 3.3: Differences in the mean state are quite hard to distinguish given the readability issues of Fig. 12.
- Lines 328-330: "In summary, the performance of the mean state simulations does not change significantly when using the modified scheme (EXP_COR), indicating that the

subgrid parameterization scheme can be incorporated into the GCMs without heavy retuning." This may be a bit overstated. It's important to acknowledge that the current analysis is dependent only on annual, global/zonal means, and only a handful of variables. Spatial/seasonal maps could strengthen the arguments made.

- Lines 382-383: "The mean states did not change much after the introduction of the new parameterization scheme, and thus, the new scheme can be implemented in the CESM without heavy retuning." In addition to the concerns raised above, it's important to note that these experiments are using an outdated version of CESM. In order to be incorporated into the model, these results would need to be reassessed using CESM2; some explanation of why CESM1.2 vs. CESM2 was used could be helpful in this case, and/or whether there are potential challenges to using this within CESM2 vs. CESM1.2.

Minor comments

- Lines 34-36: "This results in most of the precipitation simulation errors in GCMs, such as…" Are the authors saying that the use of grid mean rather than subgrid surface fluxes is responsible for the majority of precipitation biases in GCMs? That may be an exaggeration; the studies cited in the rest of the second sentence seem to suggest a range of other ways to mitigate precipitation biases, in fact. Suggest rephrasing for clarity.
- Line 36: Suggest additional citations on the bias of the rainfall intensity spectrum, as this is a frequently recognized bias that was recognized well before 2021 – the historical context is useful.
- Line 42: "…changes in vegetation density have been found to favor the release of latent heat…" Please be specific in what direction the changes are occurring; does an increase or a decrease in vegetation density link to favored LHFLX?
- Fig 1: Does this spatial pattern of correlation coefficients change with season? It would likely be more useful to the reader to consider DJF/JJA means at least to understand the importance of the correlation coefficient being a time-varying quantity rather than one that is static in time.
- Line 123: Please clarify the time scale at which the correlation coefficient r varies; is this computed at each time step?
- Section 2.3: Are the seasonal averages based on monthly means or higher temporal resolution? Daily data may yield a more realistic picture of model performance, even though it is averaged up to the seasonal level in analysis.
- Figure 2: Since the main region of interest in the TP, it would be helpful to also give the RMSE and correlation for that specific region (as it seems may be the norm in other figures of this paper).
- Figure 3 (and other figures): Suggest adding labels that denote "EXP – CTL" and "EXP_COR – CTL" to each subpanel, as in Fig 2.
- Lines 196-200: This section seeks to draw comparisons to the EXP and EXP_COR simulations, but there are no EXP-related plots in Fig. 5; please reproduce figures from *Sun et al*. that are key to this analysis within the current paper.

- Figure 5: In addition to including plots of the EXP case, it would be helpful to also include the raw EXP_COR analysis and (space permitting for this) the difference from MERRA2 observations and not just the CLT case. This would make the analysis points that are made much more convincing and clear.
- Line 211: As noted above, it does not appear that the atmospheric circulation is "corrected" in this case, so I suggest softening the language.
- Figure 6 (and others): Please clarify if the RMSE and COR values are *just* for the inset region over the TP? It would be helpful in general to have these values for both the full domain *and* the TP region itself.
    - It would also result in a cleaner plot to simply place a rectangle around the TP region where those values are calculated (perhaps in panel a) rather than an additional inset plot since the TP region is not any larger than in the main subpanel.
- Line 269: "It should be noted that there is a significant negative band along 60˚N." More explanation should be included if this is to be noted – why is this band present in EXP_COR but not EXP, does this suggest better/worse agreement with obs, etc.
- Fig 10: In this figure and similar ones that follow, it would be useful to show the raw CTL case values as well so that the reader can determine how large each change is relative to the baseline. It would also again be useful to include "EXP-CTL" and "EXP_COR-CTL" in the plot titles.
- Fig. 13: It would be helpful to show the difference in bias (relative to obs) for each case, not just the difference in each case relative to CTL. That would make the statements in section 3.3 more clearly supported.
- Lines 369-372: The precipitation improvements from EXP are *mostly* present in EXP_COR, but Fig. 2f suggests a reintroduction of a dry bias over eastern China that had been removed in EXP, which should be noted. The overestimates in precipitation along the southern/eastern borders of the TP are also still present and of a large amplitude, so with the current colorbar it is hard to deduce that those biases are "significantly alleviated."