Checkerboard Patterns in E3SMv2 and E3SM-MMFv2

Hannah et al.

We appreciate the comments from the reviewers as they have made some important points that have been addressed in the revised manuscript. This includes a new Figure in response to a comment from reviewer #2 that shows pattern examples in daily mean data from E3SM-MMF that correspond to the example patterns in Table 1 (see Figure 4 in revised manuscript).

In reviewing our pattern detection code we realized an error in the text in regards to how binary values were determined. Instead of using the neighborhood mean as a baseline to define if adjacent neighbor values are relatively high or low, we actually use the value of the center cell. This choice of what difference to use does not affect the conclusions. We have corrected the text to reflect this.

Also, the pre-print describing the CRM variance transport method to address the checkerboard in E3SM-MMF was recently posted for public discussion. We have modified the last paragraph of the current manuscript to include this citation.

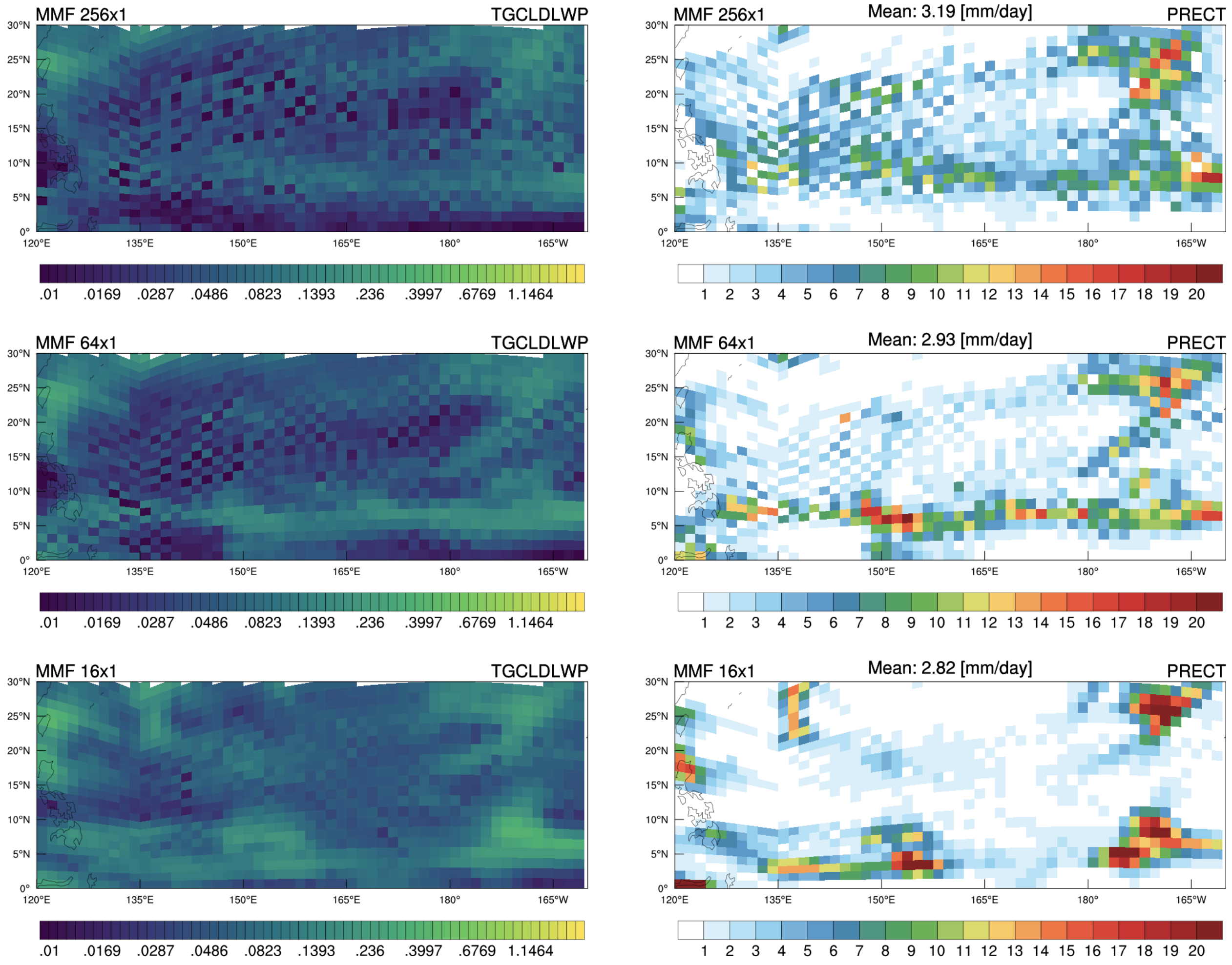
Response to Reviewer 1

A general question for curiosity: Have you performed any sensitivity tests on various resolution setups? It would be interesting to see whether the patterns occurs at all or whether they are more/less frequent under different resolutions.

We have done numerous sensitivity tests, but our primary focus was testing parameters that affect the physics behavior rather than testing the resolution of the global mesh. Unfortunately, the majority of this data was generated prior to developing a method to adequately quantify the checkerboard occurrence, so we could only rely on our eyes to judge if the checkerboard was better or worse. That data has since been scrubbed and the model has changed since then, so we would need to make a new ensemble to do a proper, systematic investigation of the sensitivity. Luckily, the variance transport scheme that is documented in the recently posted manuscript is effective at addressing the checkerboard pattern issue, so exploring the sensitivity more is less of a priority.

During the early exploration period we created some internal documentation of these sensitivity tests, and although there were a few cases where the amount of noise changed, there was a clear consensus that no amount of tinkering with parameters was going to solve this problem for us. The website where we posted plots and discussion is locked from public access, but we've shared a few plots below to satisfy the reviewer's curiosity.

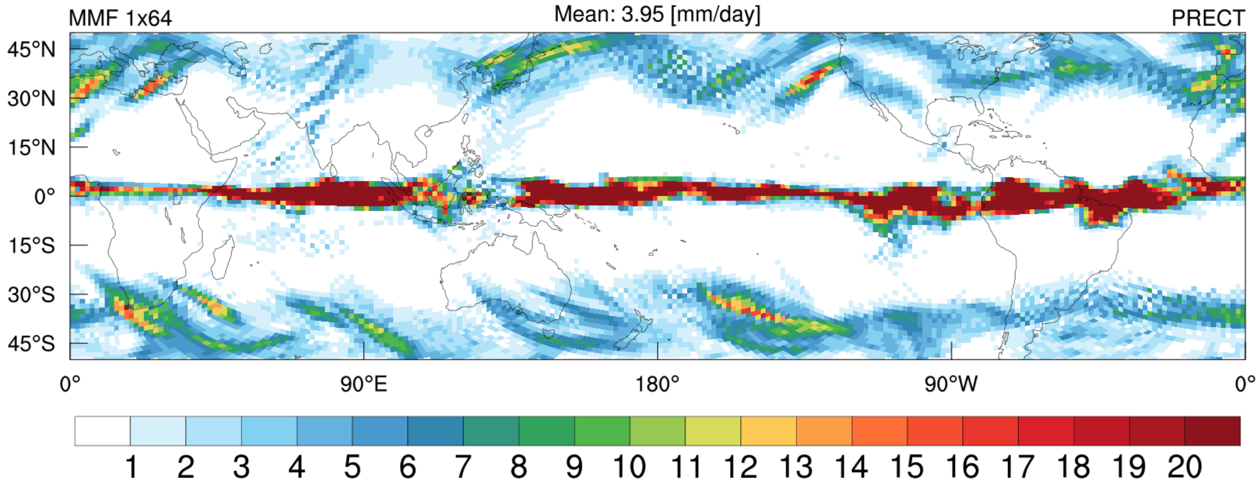
This first plot shows the mean liquid water path and precipitation from 5-day runs, in which we changed the number of CRM columns without changing the grid spacing. It's clear that a larger CRM domain produces a stronger checkerboard. At this point we were fumbling around in the dark, but in retrospect we think this sensitivity is the result of two competing factors: (1) reduced “convective throttling” in the larger domain leads to more persistent effects of variance trapping, and (2) the larger CRM domain can trap larger scales of variance (i.e. more degrees of freedom) that leads to a more obvious checkerboard signal. Note that the convective throttling causes a problematic widening of the precipitation distribution for small domains such that extreme precipitation rates become much more common, which degrades the realism of the model solution, so a larger CRM domain is always preferred.

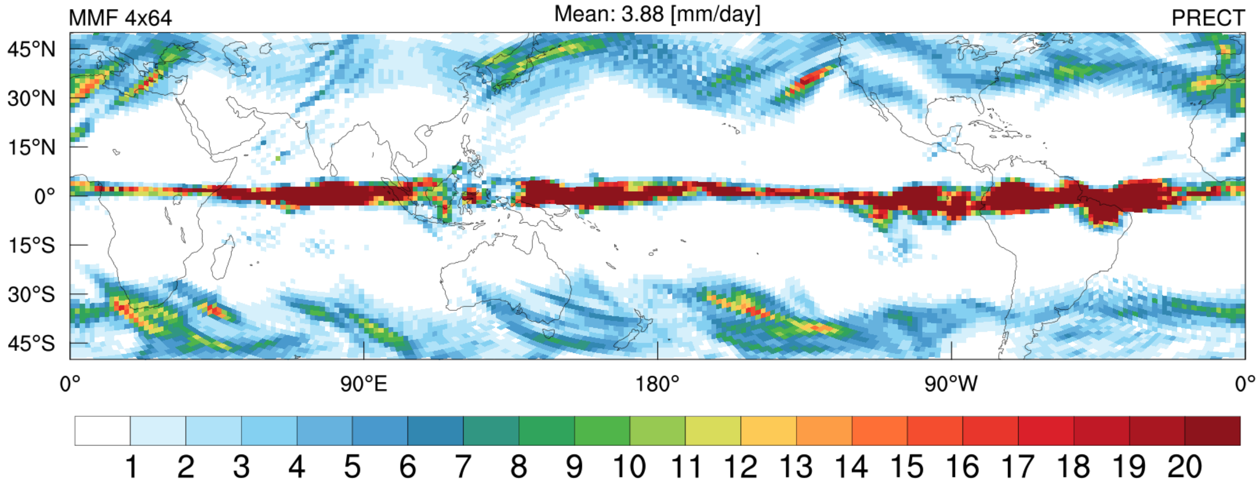


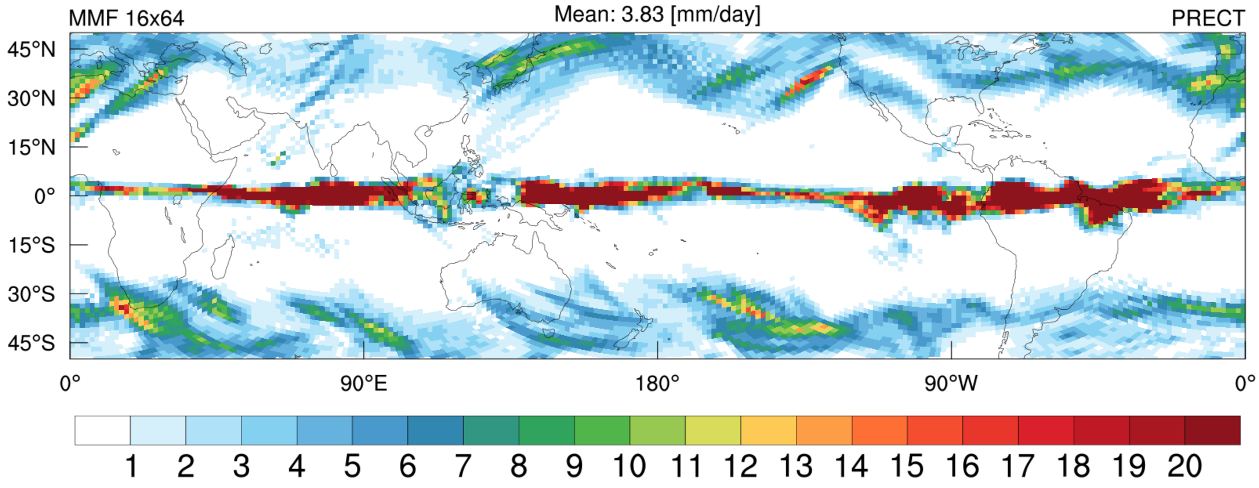
Also note that since we never ran these tests through the detection algorithm we don’t actually know if the occurrence of checkerboard is different. It might be the case that the occurrence is the same but the effective “amplitude” of the checkerboard patterns is different. Presumably we could tweak our method to provide a measure of checkerboard amplitude, but it’s not immediately obvious how this could be achieved.

This next plot compares a 4-day average of precipitation in an aquaplanet configuration using a 2D domain and 2 rectangular 3D domains, all with a north-south orientation of the longest dimension. This was meant to address one of our early hypotheses that the checkerboard might be related to the tendency for upscale energy cascades in 2D models. Thus, it seemed reasonable to speculate that a 3D CRM domain would naturally solve the issue by producing a realistic downscale energy cascade. The 3D domain does appear to help somewhat, but does not completely resolve the issue.

In this same vein we also tried adding a Newtonian damping such that the 2D CRM state was always relaxed to the horizontal mean at each level. Our thinking was that this would crudely represent a natural sink of variance by large-scale dynamics that is missing in the MMF. In both cases, our experiments convinced us that the upscale energy cascade from a 2D domain was not important in producing the checkerboard patterns.

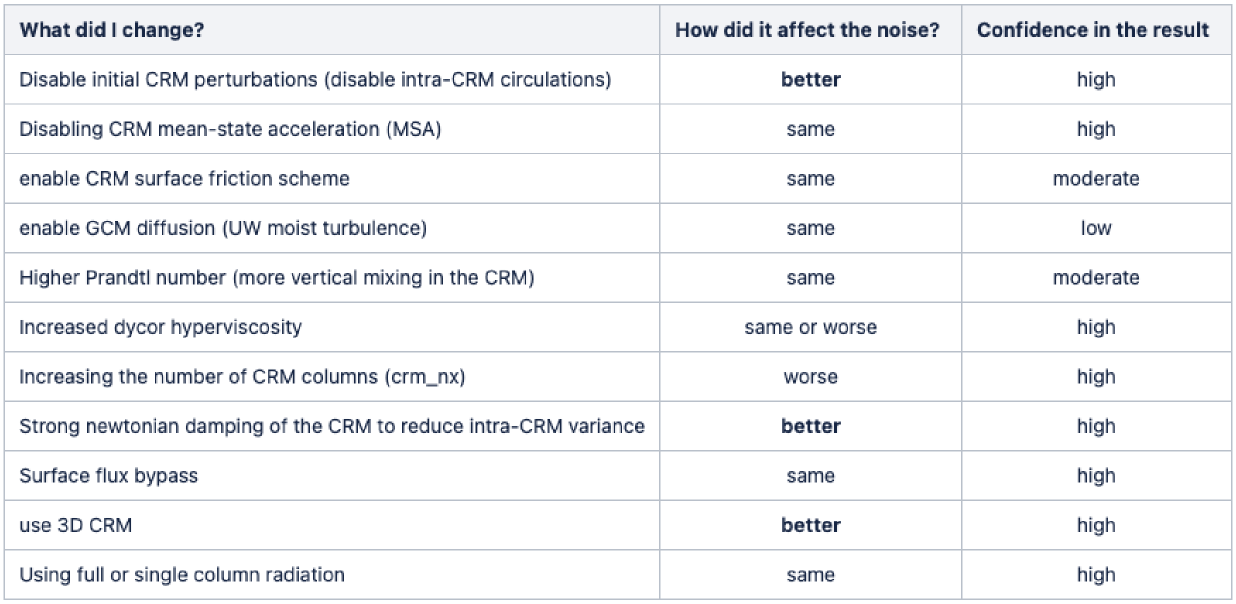






We also spent a lot of effort verifying that the hyperviscosity used in the global dynamics was not affecting the checkerboard. The simple experiment of using more hyperviscosity made the checkerboard patterns slightly worse, which surprised us since we naively expected it to help smooth things out.

Lastly, here’s a table we made during our early explorations to summarize our observations on the various sensitivity tests we had done at that point:



What did I change?	How did it affect the noise?	Confidence in the result
Disable initial CRM perturbations (disable intra-CRM circulations)	**better**	high
Disabling CRM mean-state acceleration (MSA)	same	high
enable CRM surface friction scheme	same	moderate
enable GCM diffusion (UW moist turbulence)	same	low
Higher Prandtl number (more vertical mixing in the CRM)	same	moderate
Increased dycor hyperviscosity	same or worse	high
Increasing the number of CRM columns (crm_nx)	worse	high
Strong newtonian damping of the CRM to reduce intra-CRM variance	**better**	high
Surface flux bypass	same	high
use 3D CRM	**better**	high
Using full or single column radiation	same	high

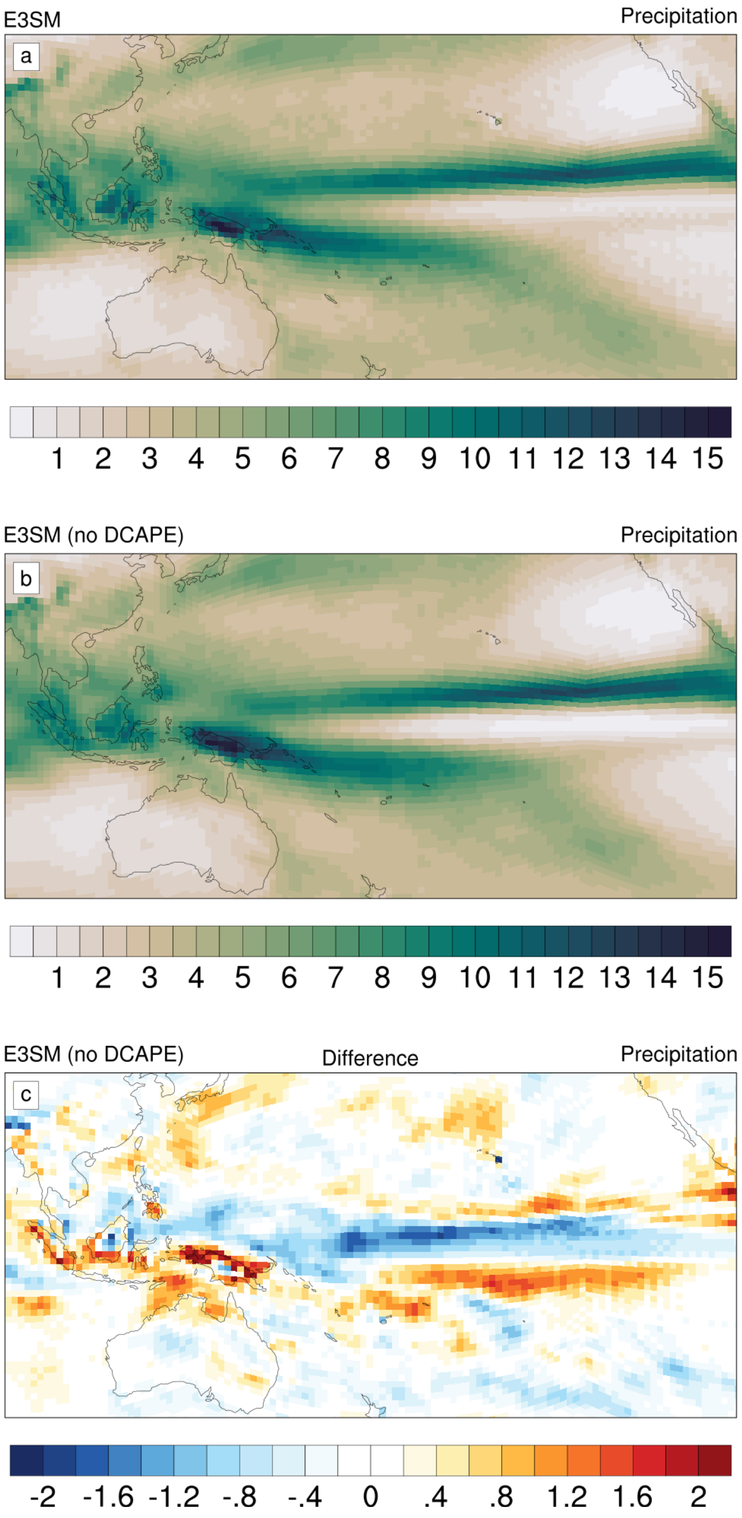
Minor comments:

1. Lines 182 - 186 and Figs 4-5: As the authors pointed out, the E3SMv2 shows less obvious patterns. It will be better to show 1) a map of difference between with/without DCAPE, and 2) maps of the differences between simulations and satellite regrided results.

We considered adding the difference maps suggested by the reviewer, but we feel that they are not valuable for discussing the checkerboard patterns. The main features in these difference maps are the large-scale climatological differences, which are not a focus of our analysis.

We also thought that the difference between E3SMv2 and E3SMv2+DCAPE would be potentially valuable to highlight the checkerboard patterns that are barely visible in the mean of E3SMv2. However, this difference map (shown below) exhibits the same problem of being visually dominated by climatological differences. The difference map also does not provide a better visualization of the checkerboard.

For these reasons we prefer to omit the difference maps from the manuscript to keep the discussion focused and concise.



2. Line 196: I believe the author meant to refer to Figure 6b,d here.

Fixed.

3. Figure 6: Is there an explanation why the simulation without DCAPE is even smoother than satellite results?

From our personal experience, ZM typically contributes to the classic “drizzle problem” in which the deep convection scheme is firing too often, but in many cases the only plumes that are allowed are the deepest ones with the smallest entrainment rate, and with a small amount of CAPE the closure restricts these plumes to have weak mass fluxes and weak precipitation.

The ZM closure is tied to the presence of CAPE and it will always fire as long as other threshold criteria are met. The DCAPE trigger further restricts convection because the mere presence of CAPE is not enough. Instead, CAPE must have been generated by the dynamical core in order for deep convection to exist.

We believe this is the reason why the data used in Figure 6 indicates a higher occurrence of the smoothest patterns, but it is unclear how we would verify this as a causal link. Therefore, we hesitate to speculate on this subject in this manuscript. It seems appropriate for a more targeted study to investigate why the DCAPE trigger affects the smoothness of the model solution, or perhaps, more generally, whether the ZM naturally leads to an unrealistically smooth solution.

The previously published studies that highlight the DCAPE trigger never assessed the potential for unphysical grid patterns and do not discuss the resulting change in smoothness as far as we are aware. This is partly because the unstructured grid data was always remapped to a lat-lon grid that masked the problem (various personal communications), but also because this sort of sensitivity is generally not expected in these types of studies that experiment with modified closure assumptions or trigger conditions in existing convective parameterizations.

Response to Reviewer 2

This article describes the checkerboard pattern that occurs in the E3SM atmospheric component in some of the cloud components (e.g., precipitation and liquid water path). This phenomenon only occurs over water and in the vertical levels where shallow clouds are present. Moreover, this phenomenon occurs in both E3SMv2 (which uses standard subgrid-scale (SGS) parameterizations) and E3SM-MMFv2 (which replaces the standard SGS parameterization with what had been previously referred to as “super-parameterization”). This phenomenon can be essentially removed in E3SMv2 when the DCAPE trigger is disabled; this is quite puzzling to me but, without understanding the mechanism of their DCAPE it is difficult to conjecture why this would happen. To capture this checkerboard pattern the authors develop an interesting technique whereby they divide the domain into groups of element (9), compute the mean and then mark the 8 elements (excluding the center element) with either a 0 or a 1 depending on whether the local mean is below or above the group mean. Using this approach, they can now easily identify the checkerboard noise signal.
The article is well-written (with only a few grammatical errors), and it does a good job of explaining

what they did although they can do better (see Minor Comments below). The authors identify some possible fixes so am wondering why they didn't address the fixes in this paper. Why wait for another paper? I would find that paper a far more interesting paper since it (1) identifies an issue in moist models, (2) presents a means to capture this signal, and (3) fixes the issue. The current paper only addresses the first two items but it's up to the authors what they want to include in this paper.

We appreciate the reviewer's thoughtful comments and attention to detail. We initially considered combining this paper with a paper on our "CRM variance transport" solution, but we felt there was a natural separation between the two topics. The pattern detection method took a lot of work to develop and is complicated to describe, so it seemed appropriate to keep the current manuscript focused on these details and discuss what we could learn from the method. Similarly, the CRM variance transport is a bit complicated to describe and it made sense to have a paper dedicated to validating that the method does what we expect. Hopefully we can write another paper on addressing the checkerboard in E3SMv2 in the future, but currently we do not have an adequate solution.

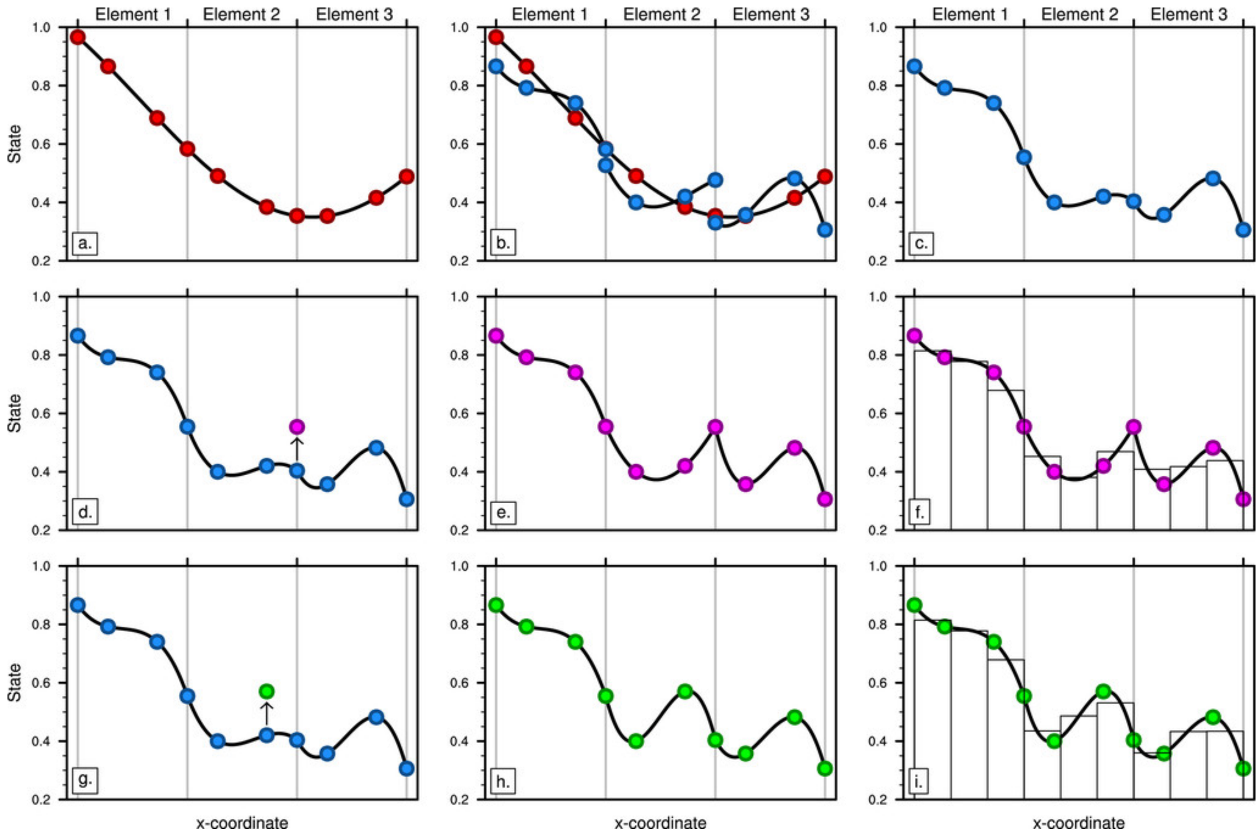
The current version of the paper would have a stronger impact on the community if more details were included that would allow for more easily understand reproducibility. I am referring specifically to their MMF setup. Even though it has been described elsewhere, it would be best to give more details here (see Minor Comments below).

(See response below)

Line 30: Why would the fact that the derivatives are discontinuous at element edges be problematic for the MMF model. The DSS operator takes care of this "discontinuity".

The DSS operator ensures that the basis functions that describe the state are continuous by averaging the element edge values shared by multiple elements. But this does not ensure that derivatives of those basis functions are continuous across a shared edge.

We would like to stress that we are not SE experts, but our current understanding is that the grid imprinting in E3SM-MMFv1 is a result of "cusps" at element edges that are a product of this derivative continuity issue as described in Herrington et al (2019) and depicted in their Figure 2e shown below.



A one-dimensional schematic showing the relationship among the basis functions, the quadrature nodes, and the proposed physics grid over the course of a time step. The filled circles are the GLL quadrature points in each element, which are connected by a Lagrange polynomials basis (curves). (a) Smooth initial conditions are (b) advanced by the dynamics one Runge–Kutta step (blue); (c) the solution after applying the DSS operator. Applying (d) gridscale forcing to an element boundary node, (e) the basis representation is clearly at the element boundary. In contrast, (g) applying gridscale forcing to an interior node (h) results in a smooth, continuous field. (f),(i) Vertical bars pertain to the values on the physics grid, found through integrating the basis functions over the control volumes. (from Herrington et al., 2019)

Herrington, A. R., Lauritzen, P. H., Taylor, M. A., Goldhaber, S., Eaton, B. E., Bacmeister, J. T., Reed, K. A., & Ullrich, P. A. (2019). Physics–Dynamics Coupling with Element-Based High-Order Galerkin Methods: Quasi-Equal-Area Physics Grid, Monthly Weather Review, 147(1), 69-84.

This feature of the SE dycor appears to drive localized convergent circulations at the element edges, albeit weak. Many of the E3SM dycor specialists have pushed back on this point because even when cusps form they shouldn't stick around for long, and will ultimately average out over time. However, we know that this quirk of the SE implementation in E3SM causes extra “noise” along element boundaries that shows up in the vertical velocity field and E3SM-MMF is very sensitive to this noise when the physics columns are colocated with the GLL nodes.

Traditional convective parameterizations, like ZM, are insensitive to this noise. We believe this is because the signal tends to be localized in the vertical, rather than producing a coherent effect through an entire column that would affect the CAPE and trigger a coherent response from the parameterized convection.

Despite being confident of these observations, we have not found a way to test these ideas directly without sinking time into designing special model configurations that have no use outside of these tests. Another approach to indirectly test our understanding is to ensure that derivatives are continuous across elements, but implementing a method for this would also be a costly endeavor that dycor experts have warned would likely not increase the accuracy of the model solution enough to justify the higher computational cost.

We have modified this sentence as follows to hopefully provide some clarity on this question without diving into all the details:

> *This problem is hypothesized to be related to "cusps" in the solution that can form due to discontinuous derivatives at the shared spectral element edges. These occasional cusps lead to noise in the vertical velocity field (Herringtonet al, 2019), which the embedded CRMs in E3SM-MMF are notably more sensitive to compared to traditional convective parameterizations.*

Line 35: If each spectral element is decomposed into 4 finite volume cells then you will have 4N_e MMF models (where N_e are the number of spectral elements). Is this correct?

That is correct. The number of CRMs used for the MMF is equal to the number of columns on the physics grid. The papers cited in this paragraph go into more detail about the differences between these grid configurations.

Line 80: Section 2.1 needs more details. This is where your new algorithm is described but I had a difficult time following it. E.g., are you doing this analysis to the spectral element (SE) cells, where you compute the dynamics, or to the finite volume (FV) cells, where you compute the physics? A picture here would go a long way to better explaining this.

Generally, we only analyze data on the physics grid because this is where all the interesting cloud stuff happens and it is what is used for coupling to the surface components, but more importantly, the physics grid is what is used for most of the model's output data. E3SM has very limited capabilities for writing out data on the dynamics grid, and none of this data is written by default (aside from restart files). We have reworded this section to clarify the adjacent neighbor identification.

Line 82: Fix near "between"

Fixed.

Line 83: Fix near “northernmost”

Fixed.

Table 1: A figure showing what these look like would be awesome.

We have added the figure below to show examples in the MMF data that correspond to the examples in Table 1.

Fixed.

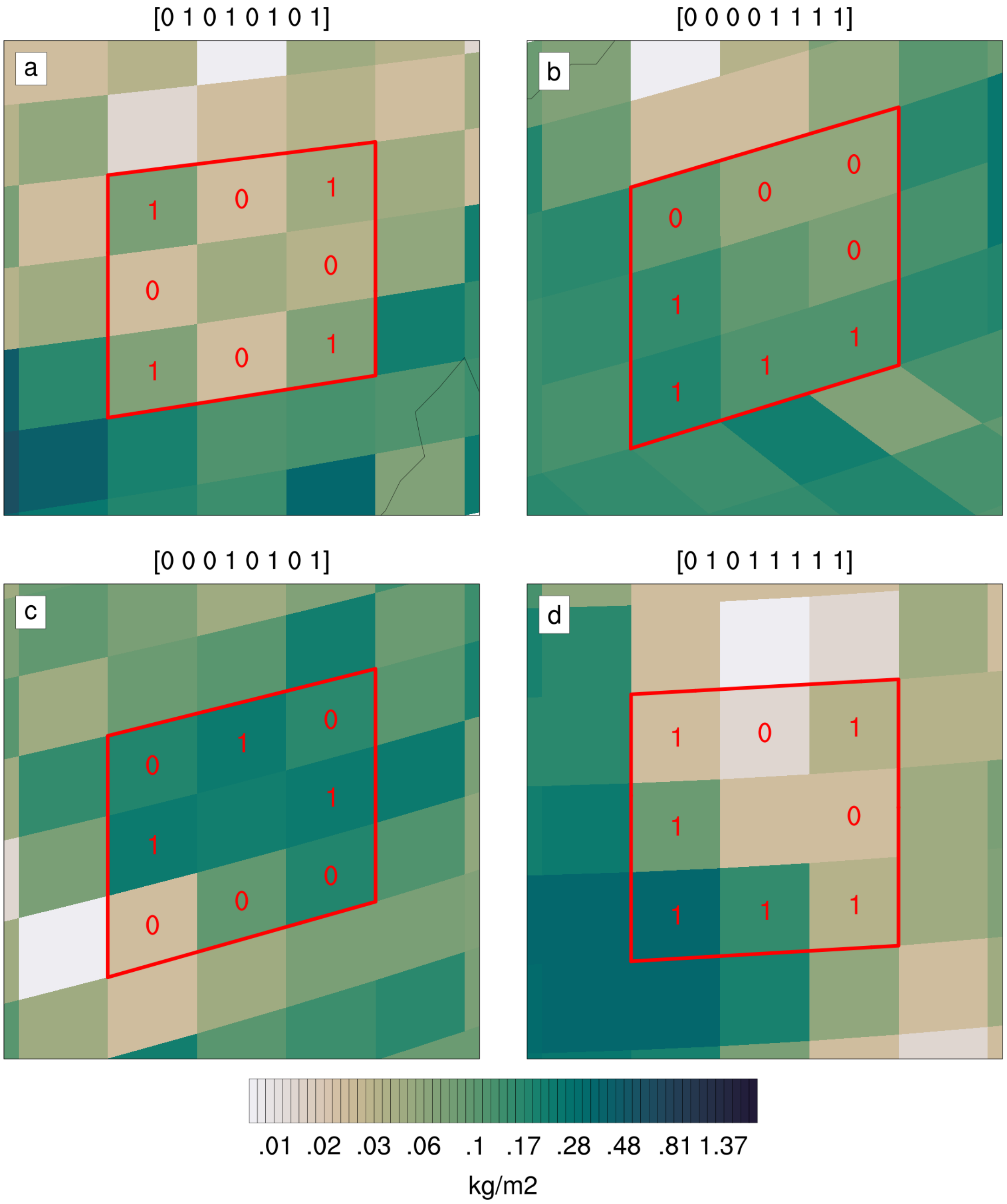


Figure 4. Examples of the patterns identified in daily snapshots of liquid water path from E3SM-MMF corresponding to the pattern examples in Table 1. Cells are labeled with a "1" for values greater than or equal to the center value and "0" otherwise. Note that a logarithmic spacing is used for the color levels.

Line 92: Fix near "equal"

Fixed.

Line 93: Why ignore the center cell?

As noted at the top of our response, we discovered that we had made an error in the description of our pattern detection method. Instead of using differences from the neighborhood mean to encode the binary values (which was our initial approach) we use differences between neighbor values and the value of the center cell.

But if we had used the neighborhood mean to define these differences, then including the center point in the pattern detection complicates the implementation and interpretation, although we do not think it would affect the general results. In this scenario, we would have to give the center value an arbitrary position in the binary sequence, which complicates how we handle the rotational symmetry. This also makes understanding the distribution of other patterns much more difficult (ex. Figure 6) because the identification of local extrema is more complicated. Also, the process of identifying partial checkerboard instances becomes much trickier since we can't consider alternating sequences of 1's and 0's and have to consider the center value in any sequence of neighbor values. We have slightly reworded this paragraph to clarify this point.

Line 127: A clear description of the CRM is essential here. E.g., how wide is the horizontal, what are its dimensions, etc? Is it too small so that the periodicity keeps things localized and moist components remain captured? Is the MMF domain 2D, 3D? A brief, but clear, description of your setup is imperative.

These details were included in the last sentence of this paragraph starting on Line 129 of the originally submitted version:

> "*The embedded CRM in E3SM-MMF uses a two dimensional domain with 64 CRM columns in a north-south orientation and 1 km horizontal grid spacing.*"

To clarify this concern, the size of the CRM domain does not influence whether the cloud-scale fluctuations are trapped or not, but small CRM domains do cause other problems (see work by Mike Pritchard et al. on the "convective throttling" issue) which appear to influence the checkerboard patterns. In our reply to Reviewer 1 above we described some of our initial sensitivity tests, which largely found that no amount of tinkering with the configuration made the checkerboard patterns disappear. Even in our followup paper about the CRM variance transport method we found that despite being able to eliminate the checkerboard in the climatology, we could not completely reduce the checkerboard pattern occurrence to levels comparable with satellite data on short time scales. We have noted in the revised manuscript that these configuration details do not qualitatively affect our results.

Line 143: Sec. 2.4 contains your model simulations. More details would be helpful here. E.g., do you only use MPI? MPI with OpenMP for multi-threading on KNL hardware?

The CRM itself cannot utilize CPU threading and the individual CRM instances are not decomposed with MPI despite having this capability in the standalone SAM maintained by Marat

Khairoutdinov. Luckily, we can utilize threads in other parts of the atmosphere physics, which helps out a lot in the performance of radiation, but in these simulations threading was not used because we find that we don’t get a noticeable benefit on machines like Cori-KNL. We already mention that we use 5400 PI ranks, and we have added a sentence to clarify that no threads were used.

Line 198: fix near “occurrence”

Fixed.

Line 285: fix near “balance”

Fixed.

Line 304: is the rest of the physics the same? Is this why you think it is the dynamical core? Doesn't E3SM-atmos use nonhydrostatic equations while CAM uses hydrostatic? Any other big differences you can identify?

In our comparison with SP-CAM, both models were configured to use the hydrostatic dycore. We’ve looked into many ways of ensuring that we are performing a fair comparison between the two models. While some differences are easily remedied (ex. Number of vertical levels and CRM columns), and others can be compensated for with a bit of work (ex. radiation scheme), other differences are unavoidable, such as the choice to use the full pressure in the CAM-SE dycor instead of the “dry pressure” that is used in the E3SM dycor. We’ve intentionally chosen to avoid describing these comparisons in detail because we don’t fully understand how they could be so different. We would like to probe this question more, but given the likelihood that SP-CAM will cease to be supported in the near future we may never know the explanation behind this difference.

Line 311: Regarding your conjecture that the “scale gap” is a design flaw of the MMF approach. Perhaps the issue is that the GCM is not imposing as big an influence on the CRM and, therefore, the CRM dynamics remain trapped (but I suppose this will always be the case when the CRM is allowed to evolve independently, no? Perhaps this is what you are saying). So the fix is to let these processes out - your variance transport. Would be very interested in hearing about how this is applied and could very well save the MMF idea.

As discussed above, we chose to put the variance transport description in a separate paper that was submitted after the present manuscript. To spoil the surprise, the general results is that the model climate is not significantly affected by the variance transport scheme, other than strongly reducing the checkerboard pattern occurrence. We did observe a reduction in temporal variance of certain variables on the GCM grid, which is a positive development since the default E3SM-MMF configuration produces too much variance.

As a side note, it's tempting to think about the MMF as one model having a strong influence over the other, but this is not an accurate portrayal. The GCM and CRM are coupled at a specific scale such that the localized GCM state and the CRM domain mean are exactly the same. So it's better to think about the coupled system of two models in which both models have equal influence over the state at a given location. By design, this coupling approach ignores fluctuations from the CRM horizontal mean. There's no clear way to formulate the MMF coupling scheme to provide a natural interaction between these fluctuations and the GCM, unless these fluctuations were to occupy a specific location within the parent GCM column, rather than vaguely existing as a "representative sample" of the GCM's unresolved sub-grid structure.

Figure 1 Caption: fix near "scale"

Fixed.