Thank you for your constructive comments, which were useful to improve our paper. Please see our responses below. The comments are in **bold** *italics* while the responses are in normal type.

I encourage the authors to provide a deeper discussion and interpretation of the results. For example, the lack of relationship between convective precipitation and near surface vertical velocity (Figure 2b, 3b) and the mismatch in timing with CAPE/ CIN in CanAM4.3 relative to spCAM5 are particularly interesting findings.

We added the following text in Section 4.4 of the manuscript:

"Therefore, a transition from a large-scale subsidence to large-scale ascent may be important in triggering convection. A near-surface omega tendency has been previously used as a trigger in the Donner convection scheme (Donner 1993; Wilcox and Donner 2007) in a version of the Geophysical Fluid Dynamic Laboratory (GFDL) Atmospheric model, version 3 (AM3) GCM. In their model, convection is triggered when near-surface omega becomes positive and exceeds a specified value and convective inhibition is less than 100 J kg-1".

"Convective precipitation in CanAM4.3 does not seem to correlate well with CIN (Fig. 3b) and this is likely because CIN is not independently included in the ZM closure in CanAM4.3. Therefore, any discussion of CIN and linkage to CanAM4.3 precipitation is out of the scope of this paper. We should point out thought that, CIN is tightly coupled with precipitation over mid-latitude summertime continent but not with precipitation over oceans (Myoung and Nielsen-Gammon, 2010)."

How are the deficiencies in the parameterizations used in CanAM4.3 (i.e., CAPE based closure), which have been identified here, different from what is already known and published?

In the introduction we briefly mentioned that some commonly used convective scheme in GCMs employ triggers and closures based on convective available potential energy (CAPE) or CAPE generation while other closures are based on net column moisture convergence. Other convective schemes, for instance the Donner convective scheme, use grid-scale upward motion in the lower troposphere as trigger function. Although, the Zhang-McFarlane (ZM) convection scheme is very popular and has been modified and improved over time, as described for example in Zhang and Mu, 2005a, the ZM scheme still has deficiencies, such as, generates too frequent too light precipitation and underestimates the frequency of extreme events. However, various models employ various version of the ZM scheme and our goal is not to modify the ZM scheme employed in CanAM4.3, but to compare the precipitation generated within the ZM scheme to precipitation generated within a cloud-resolving model under similar large-scale forcings.

And how can new information from the results presented here be applied to further improve models beyond what has already been implemented?

One new result is that precipitation generated within a cloud-resolving model depends on both CAPE generation and near-surface omega, two commonly used variables in the trigger and closure functions of most popular convective parameterization schemes, while convective precipitation is a function of CAPE only in a CAPE based closure model (CanAM4.3). Another new result is that, the cause-consequence analysis (Figure 3) show that variations in omega precede variations in convective precipitation while variations in CAPE generation trail variations in convective precipitation in spCAM5. Based on these results we suggested that near-surface omega might qualify as better trigger and may be used together with dCAPELSFT in the closure scheme in CanAM4.3.

The effort to calculate "convective precipitation in spCAM5" in a way that is comparable to "convective precipitation in a parameterized model" is a great idea and potentially very useful. However, it is not clear that the way it is calculated in spCAM5 here means the same thing as it does from parameterized convection in CanAM4.3. How sensitive are the results to the values of the criteria (vertical velocity and cloud water/ice)? More importantly, how well does a definition of "convective precipitation " based on CRM vertical velocity and cloud water/ice match what "convective precipitation" means in a global parameterized model? Since the comparison and analysis is contingent on this calculation, it would be useful to discuss other ways it could be defined within spCAM5 and/or expected differences with what convective precipitation means in CanAM4.3. It would also be helpful to use an independent calculation of "convective precipitation" that could be applied identically to both models, which would likely be dependent on large-scale conditions. Ultimately, to what degree do the results and comparison between the models depend on the way that convective precipitation has been defined? Likewise, how is CAPE calculated in spCAM5, is it at the CSRM or GCM scale? A comparison to CAPE calculated at the GCM scale would be most consistent with CAPE from CanAM4.3. Along these same lines, the differences in the relationship of convective precipitation and omega between spCAM5 (strong correlation) and CanAM4.3 (no correlation) may be, in part, due to differences in the definition of convective precipitation. I suggest including some analysis of relationships with "total precipitation rates" or alternative definition of "convective precipitation" in spCAM5.

We agree with the reviewer that the definition of convective precipitation in spCAM5 will be sensitive to the values of the vertical velocity and the cloud water and ice. The method we used to define convective precipitation follows that in Suhas and Zhang (2015) and Song and Zhang (2018). Using this method, 68 % of the total precipitation in spCAM5 was convective compared to 71 % in CanAM4.3.

However, as the reviewer suggested, we further investigated the sensitivity of our results to the definition of convective precipitation. We repeated all the analyses using total precipitation instead of convective precipitation and generated Additional Figure 1, 2(a), 2(c), and 3 (below). The results in the Additional Figures are similar to those in Figure 1, 2(a), 2(c), and 3 in the manuscript. Therefore, we can say that the findings in the manuscript are not sensitive on the details of how the rainfall is partitioned in both spCAM5 and CanAM4.3.

We added the following text in Section 3.1 of the manuscript:

"The sensitivity of the results to the definition of convective precipitation from spCAM5 was evaluated by repeating the analyses using total instead of convective precipitation. The results in Figure 1, 2(a), 2(c), and 3 were found to be similar using either the total or convective precipitation from spCAM5, implying insensitivity, for this study, to the exact definition of thresholds in the method of Suhas and Zhang (2015)."

In general, an explicit inclusion of observations for comparison would be helpful to the reader. The authors note that there is no dependence of convective precipitation with CAPE in spCAM5, which they say is consistent with observations by citing Mitovski and Folkins [2014]. It would be useful to make this calculation and include the observations in the figure for both CAPE and dCAPE. Likewise, the authors note that spCAM5's relationship between min/max CAPE and the timing of rainfall is consistent with observations by referring to Mitovski and Folkins [2014], but again I think showing the actual observations (as referenced) on the same figure would help.

Mitovski and Folkins 2014 used 12-hour vertical profiles of temperature and specific humidity to compute CAPE. In addition, they used 3-hour TRMM 3B42 rainfall to isolate rainfall events. In this paper, however, we use sub-hourly model data to compute CAPE and investigate the relation with convective precipitation. Although the temporal resolution of the data used in Mitovski and Folkins (2014) and in this paper is different, it has been previously shown that tropical convection exhibits similar behavior on various time-scales (Mapes et al., 2006: "The mesoscale convection life cycle: Building block or prototype for large-scale tropical waves?"). The similarity in the observed and simulated (spCAM5) CAPE variation, once again shows that, in absence of higher-resolution observations, spCAM5 may be useful in studying convection-large-scale environment interactions.

To make it clear that Mitovski and Folkins 2014 use 12-hour soundings, we added the following text in Section 4.4 of the manuscript:

"12-hourly"

Minor Comments:

Why not evaluate the ZM scheme as implemented in the conventional CAM5 to have more consistency with spCAM5? Many other aspects of the model are different between CanAM4.3 and spCAM5, beyond just the representation of convection, which makes the comparison somewhat unconstrained. I suggest including results from CAM5 as well as CanAM4.3 and spCAM5. Since only 3 months of simulation time is being assessed here and the initial setup of CAM5 would be the same as sp-CAM5, this should not add a significant amount of work.

We agree that it would be interesting and useful to perform the analysis using CAM5 simulations that are configured the same as spCAM5. However, this is a non-trivial amount of work due since the spCAM5 data we used was archived from previous simulations and we no longer have access to the personnel and computer accounts. To perform the CAM5 would require significant effort to set as it would require setting up the model on a new computer system with all of the associated effort to verify it is implemented correctly. Repeating the analysis with CAM5, and potentially other model, would be something that could performed in future studies.

I am confused about the vertical resolution used in spCAM5. Typically, the vertical resolution is 30 levels in the global grid and 28 levels in the CSRM (coinciding with the lowest 28 levels). Here the authors state that there are 66 levels CAM5, which would imply 38 levels above the CSRM rather than the typical 2 levels. Have previous studies used this configuration? Have you evaluated the differences between using 30 and 66 levels Additionally, the Khairoutdinov and Randall (2001 and 2003) references are fairly old and refer to the implementation of super-parameterization in older versions of CAM. I recommend the authors cite more recent papers describing the implementation in CAM5, such as Wang et al., 2011 (https://www.geosci-model-dev.net/4/137/2011/gmd-4-137-2011.pdf).

We stated all CAM5 and CanAM4.3 levels in the atmosphere. As the reviewer suggester, we updated the manuscript and include the lower atmosphere levels only, as well as, we cite Wang et al., 2011. We substituted the following text in Section 2:

“66 vertical levels from the surface to 5.1 x 10e-6 hPa “

With:

“30 vertical levels from the surface to 3.6 hPa”

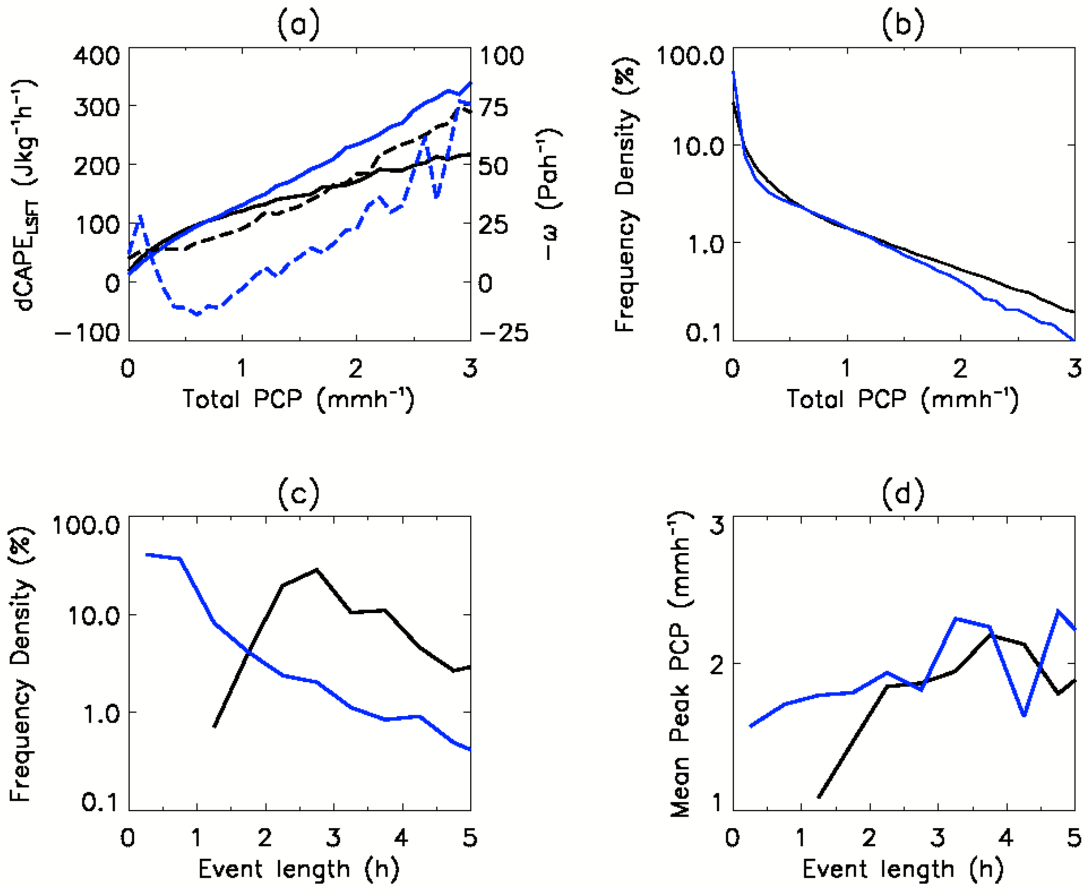
Since spCAM5 is used instead of spCAM4, it includes aerosol processes and two-moment microphysics, so it might be helpful to describe these components of the model (MAM3 aerosol and Morrison microphysics) and compare them with the same processes in CanAM4. The representation of aerosol and cloud microphysics are likely to influence precipitation as well.

We agree that it is possible that the aerosol and cloud microphysics formulation could influence the precipitation but our hypothesis is that the main control in the Tropical Western Pacific is rainfall from the deep convective scheme. In the paper we note that roughly 70% of the rain is convective and it seems that the stratiform rain has little effect on the results (performing the analysis using the total rain or the convective rain give similar results). We leave it to the interested reader to refer to the references for details about the aerosol and cloud microphysics parameterizations.

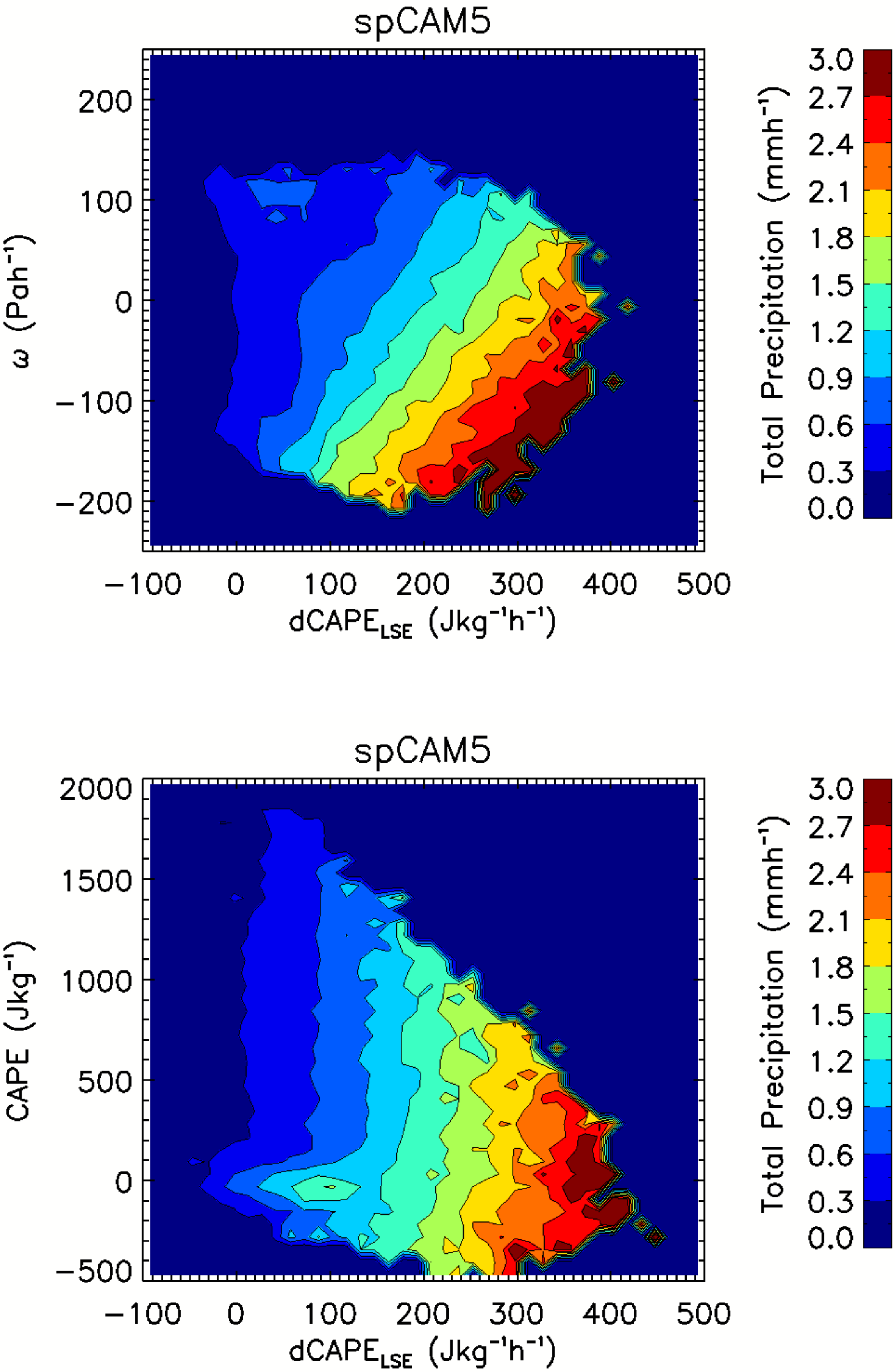
For the relationship between vertical velocity and convective precipitation in CanAM4.3 (Figure 1a), the authors conclude that "the results are not considered robust due to the few samples". Why not use more years for the CanAM4.3 results? CanAM4.3 is relatively cheap to run, so it is unnecessary for the authors to limit their analysis to such a short period. I recommend using more data, at least for CanAM4.3, to produce more robust results

We found this suggestion very useful and we therefore perform another five CanAM4.3 ensemble simulations for the period of study. The ensemble was generated by changing the random number seed on 1 January 1997. We incorporated the data from these simulations into our analysis. Figures 1, 2, and 3, are now based on 5 CanAM4.3 ensemble runs.

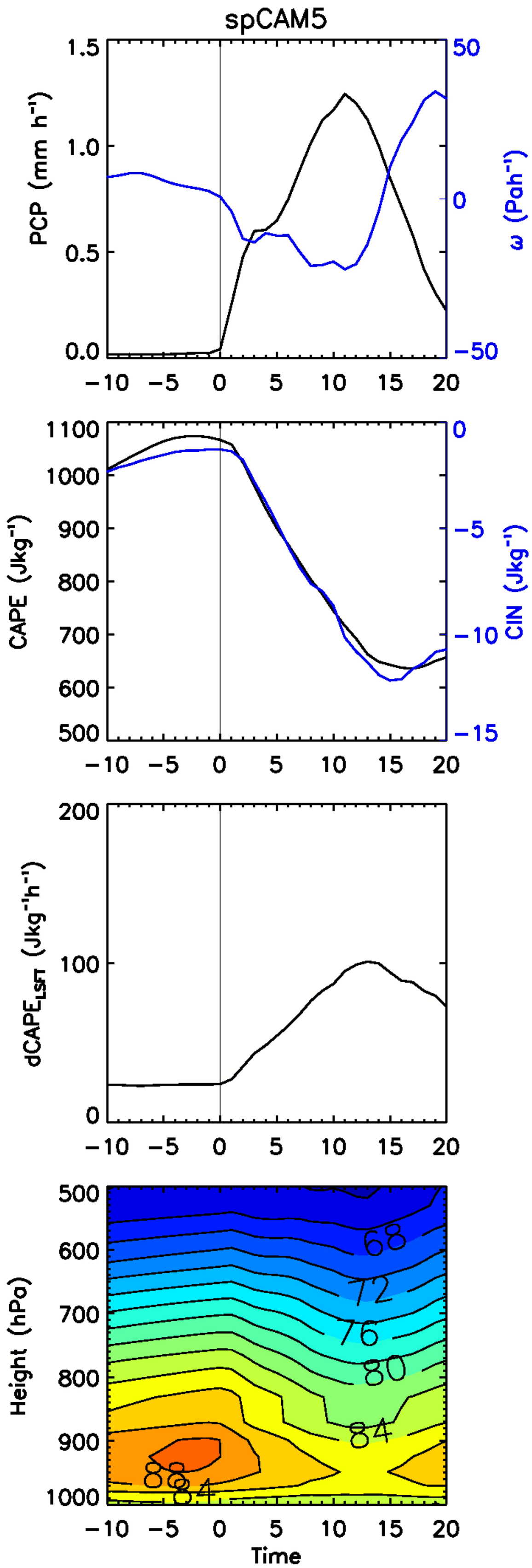
We added the following text in Section 2 of the manuscript:
“For CanAM4.3, a six member ensemble was generated by uniquely adjusting the seed for the random number generator on 1 January 1997. This was done to improve the statistical representation of the results from this model as data from all ensemble members were used in the analysis’.



Additional Figure 1



Additional Figure 2(a) and 2(c)



Additional Figure 3