Response to Reviewer #2:

We thank the reviewer for the constructive criticism that led to improvements in the manuscript. Following the suggestions from the both reviewers, we have carried out a major revision. The discussions and readability are improved and the redundancies in figures are removed. Among other things, Figs. 4, 5, 6, 8, and 11 are revised, while Figs. 9 and 10 are removed to declutter the discussion. Please note that few of the reviewer comments may not be valid anymore due to the revision of either the underlying figure or the text.

Please find below the point by point reply to your comments.

This study evaluates the performance of various versions of the EC-Earth model against satellite retrievals from CloudSat/Calipso focusing on the clouds and cloud radiative heating rates (CRH). The authors evaluate the intraseasonal and interannual variability of CRH in the tropics. This study is useful and would be a valuable contrition to the related and increasing body of literature that deals with the coupling between the circulation and CRH. However, the authors mostly illustrate a large amount of results without providing enough insights and deeper analysis of the results. Additionally, my impression was that the authors did not properly illustrate the gained knowledge from this study (potentially by showing many results without a clear story?) and did not communicate the main key points of this study to the reader. Another general remark could be the fact that the comparison among the different model versions was not giving any extra information to the paper and thus the authors should reconsider to focus on a specific model version and try to provide more deeper analysis and explanations for the differences between the EC-Earth model and the observations. As a final major remark, the readability of the vertical profiles was very bad and the authors should reconsider using more distinguishable colors.

We thank the reviewer for recognizing the potential of the results. We understand that the main message probably got cluttered in the way the analysis was presented. The other reviewer also raised similar concerns. We have revised the manuscript, considering the comments from both the reviewers, and improved the presentation, clarity and figure quality.

1) There is a confusing part in this study that comes from the fact that the authors do not clearly explain the method that they use to compare the models with the observations. Is it monthly mean, 3-hourly data that the authors use to compare the simulated CRH with the observed ones?

The model output is written out every three hours. From these results, we interpolate to the local time 13:30 to match the satellite overpass. All mean values (monthly /ENSO) are then based on these interpolated data. We have updated the paragraph starting on line 115 to “Note that the satellite simulator for evaluating heating rates in EC-Earth is currently not available. CloudSat and CALIPSO pass the equator at roughly 13:30 local time during daytime, so the model results are linearly interpolated from the two nearest output times (3–hourly data) to fit the satellite overpass time. This time interpolation together with the ability of the active satellite instruments to detect thin clouds reduces the need for satellite simulator otherwise commonly used for passive instruments (see, e.g. Pincus et al., 2012). All means (monthly/ENSO) in this study are based on these interpolated data.”

2) Another general remark is that the authors in the results section mostly describe the figures without conceptually connecting the key points of the study and without providing more insights and explanations.

We have added more discussions and potential reasoning behind the observed biases in the revised version. The main aim of the manuscript is to compare/evaluate the CRH to understand if the EC-EARTH versions are able to simulate them well to a first order and which version is closest to the observations. Therefore, we had focused on documenting these biases in CRH and the relevant cloud properties, avoiding to speculate too much on the effects of underlying parameterizations (without doing the required idealized experiments). We hope that the new additions relating CRH biases to cloud properties and other potential factors (e.g. surface parameters) would provide more information in this context.

3) The authors illustrate the changes in CRH across seasons and different ENSO phases, however they do not illustrate the gained knowledge from this changes and do not properly communicate the key message from these comparisons in the context of previous studies.

We would very much like to discuss our results in the context of previous studies. However, to our knowledge, there is no previous study that specifically evaluates the vertical structure of CRH during ENSO phases in a climate model using combined CloudSat+CALIPSO data. Almost all of the previous studies have focused on investigating the top of the atmosphere cloud forcing during ENSO in different contexts, such as understanding cloud feedbacks.
The main message from our evaluation of EC-EARTH is that all model versions capture the shift in meridional CRH associated with the changes in Walker circulation during the positive and negative phases of the ENSO. However, we see similar underestimations/overestimations in the vertical structure of CRH in the model climatology compared to the observations. This means that the basic description of cloud processes has to be improved first in the model before studying the general role of cloud radiative effects in the ENSO variability, especially related to cloud feedbacks.
We have clarified this in the revised version.

4) The readability of the figures is really bad (particularly those that include vertical profiles), thus more distinguishable colors should be used.

The figures are updated with clearer colours and larger fonts.

Minor comments:
Line 3: Typo: coupling

Corrected.

Line 148: Typo: analysis

This is now changed.

Line 159: It could be easier for the reader if the figures and the text had different longitude values to something like 45E, 90E, 90W, 45W.

Please note that we have shown plots from 0E to 360E instead of 180W-180E in order to have Pacific Ocean at the center of the plot in order to highlight the variability associated with ENSO.

Figure 2 (e-h): To make the comparison easier since the differences are larger than 1K/day wouldn't it make more sense to use the same colorbar as in panels (a-d) so that the reader can compare the differences compared to the absolute values?

We agree and have changed the colourbar in the revised manuscript.

Line 179: Do the authors refer to the temperature tendency due to convective parameterizations or the CRH? It should be properly phrased here.

It is the CRH. Clarified in the revised text.

Line 180-181: Since the model overestimates the magnitude of the CRH over both convectively active and stratocumulus regions is there a reason to separate them by using "but"?

Thanks. It is corrected.

Lines 183-184: Any potential explanation for that?

It is explained in Section 3.1.2 that discusses meridionally and zonally averaged differences. "Figure 6 shows that the models underestimate the cloud fraction as well as cloud liquid and ice water content in the lower and middle parts of the troposphere. This underestimate results in suppressed shortwave heating and longwave cooling in the middle troposphere in the models, thus, at least partly, explaining the differences observed in Figure 5."

Lines 184-186: This is confusing. Isn't it the case that panels (e-h) in Figure 2 show monthly mean differences between the model and the observations? See major comment 1

Figure 2 shows the monthly mean differences for the local time 13:30. Since the satellite only covers this time we interpolate the model results from the nearest 3-hourly data to the local time 13:30 for each day, then we calculate the monthly mean from these interpolated results. Hopefully this is clearer in the new manuscript.

Lines 189-190: It is not obvious that the convection is strong. This is not shown anywhere. Moreover, the authors do not provide enough evidence when is the dry and the wet season for a location. Maybe show this somewhere?

We have added references and updated the text in the revised manuscript.

Lines 187-199: Any further insights about these differences?

Two main factors can potentially explain this behaviour. The convection is strongly influenced by the surface temperature in the models. By improving the spatial resolution, the subgrid-scale variability in the SSTs is better represented and the modelled convection is more sensitive to the SSTs (compared to coarser grid values), leading to an even stronger hydrological cycle in the high-resolution version. Such behaviour is also seen in the regional climate models. Secondly, following the HighResMIP (High-Resolution Model Intercomparison Project) protocol, the same tuning is applied to the higher resolution version as well which may not be optimal and results in clouds that are too thick/include too much cloud water/live too long.
We have updated the discussion on this in the revised manuscript.

Lines 204-205: Isn't it the case that the magnitude in the lower troposphere (below 2 km) is smaller across the model versions, thus underestimating the magnitude of the cooling?

Yes, it is clarified in the new version.

Lines 210-212: Any insights about these differences? Could you potentially explain why the models simulate the maximum heating higher that in the observations? Moreover, the different model versions seems not to be sensitive to CRH. Is there any explanation why these differences are so small?

These differences are explained in response to the earlier comment.
There are a couple of factors that can cause the maximum heating in the models to be higher than in the observations. To begin with, the models heavily underestimate the cloud fraction between 5 to 12km altitude. These are typically deep convective and nimbostratus clouds that produce substantial cloud top cooling in the layers above (12-14 km). This cooling in the uppermost troposphere is missing in the models (Fig. 5). Instead, the clouds are predominantly located in the 12-14 km region in the models. The net effect of this is that the altitudes where observations show strong cloud top cooling are also the regions where models could have either heating or suppressed cooling, resulting in the maximum heating higher than the observations. This discussion is added in the revised manuscript.

The Primavera model is based on an early version of the model used for CMIP6 and differences between the modes are mainly in the surface albedo scheme (see Section 2.1). Therefore the large differences in the cloud radiative heating between these two different versions are not expected. However, these versions have not been evaluated before, therefore we believe it is a valuable comparison. CMIP6 version is the most state-of-the art, while PRIMAVERA versions are useful to investigate the impact of spatial resolution.

Figure 6: It might be better to use letters to show which panels show IWC, LWC and CF.

This is added in the new version.

Lines 218-220: This is somehow confusing. Despite CF differences in the upper troposphere are small, the models seems to underestimate IWC compared to the observations and and still producing stronger CRH than the observations. Could you explain why is this the case?

The underestimation of cloud fraction and cloud liquid water path in the middle and lower troposphere (Fig. 6) allows more LW radiation to reach the upper layers, where it gets trapped and absorbed. Due to lower ice water in the upper layers, the resulting LW cooling is also lower at the cloud tops. This is now explained in the revised version.

Lines 214-215: This underestimation of CRH is mainly evident in the mid-troposphere. However, in the lower troposphere CRH agree more despite the fact that cloud properties differ substantially. Any explanation regarding that?

There are also disagreements in the lower troposphere, but they are not as strong as in the middle and upper troposphere (Fig. 5). When we look into the individual SW and LW heating components of the total CRH, we see that the opposite biases in these components lead to the lower net biases in the total CRH in the lower troposphere.

Lines 244-253: Could the authors provide more insights that would explain these differences?

The nature of these differences is similar to the ones reported in Figs. 2-4. Here as well we notice that the models underestimate middle and upper tropospheric cloud fraction (between 5 and 10 km), underestimating heating and cooling in the ENSO positive and negative phases respectively. This is explained more clearly in the revised version.

Figure 7: As before, either adjust the colorbar to make the comparison more straightforward or mention in the figure caption the different colorbar for the differences. Additionally, in the figure caption it would be more appropriate to mention the satellite (CloudSat/Calipso) instead of mentioning CRH from satellite.

We have updated the colourbar and caption in the revised manuscript.

Lines 259-260: Could the authors providence more evidence why is this happening?

This comment is not applicable anymore. Figs. 9 and 10 are removed in the revised manuscript to avoid redundant information.

Line 262: Why the differences are mostly evident over the ENSO region? Is this consistent with previous modeling studies? What could be the reason behind these differences?

Due to the removal of Figs 9 and 10, this section is now revised and the differences are explained.

Line 274: Which figure shows that? Maybe add the figure at the end of the phrase?
Figure 11. This is clearer in the new manuscript.

Line 282: As before, which figure shows that?
Figure 11. This is clearer in the new manuscript.

Lines 295-298: Maybe use the same wording for anomalies? Using anomalies and changes in the same phrase might be confusing for the reader.

This is corrected in the revised manuscript.