Response to Referee #1

We thank the referee for their review. Below are the original comments in *italics* intermixed with our responses in normal text.

The paper documents the scientific structure and basic behaviors of version 5 of the Canadian Atmospheric Model. This model has and will likely continue to advance understanding climate and climate change and is a key participant in international modeling activities in support of ongoing assessment of climate change. The paper is generally clear in its explanations and presents a range of model simulations compared with observations, providing essential information for those using the model or its simulations for a range of scientific and societal impact activities.

Specific Comments and Questions
1. Regarding the cloud microphysics factors in Table 1: Are facacc, facaut, and uicefac factors that multiply nominal values for accretion rate, autoconversion efficiency, and ice fall speed? If so, what is the basis for the nominal values? The CanAM5 factors, are quite large, factors of about 10 to 6,000. Especially if these factors scale physically based parameters, is there an issue of physical plausibility? Further explanation and context would be helpful here. Related, on p. 9, l. 35, it is stated that autoconversion rates are not scaled, but the factor scaling efficiency in autoconversion (facaut) in Table 1 is listed as 0.1204. Clarification or additional information is suggested.

We agree that the original text was not clear, having mixed factors that scale the process rate (facacc and facaut) and values that are parameters in the formulation (uicefac, which has a value of 770 in the original formulation). We have adjusted the text to clarify these factors and the values.

2. pp. 12-13, ll. 33-1, Fig. 3: The "notable increase in southern hemisphere low cloud in CanAM5" relative to CanAM4 is not evident on Fig. 3.

We agree, it is not clear upon rereading the text what feature to which we were referring. This text has been removed.

3. Fig. 4: What do the solid contours on the two uppermost panels on the right indicate?

These contours are the zonal mean cloud amount from CanAM5. The caption has been updated to make this clearer,

"Zonal cloud fraction and cloud phase from CanAM5 compared with GOCCP and MODIS observations averaged over 2007-2009. Black contours on the upper and middle plots in the right column are the zonal mean cloud fraction from CanAM5. Bracketed numbers in the bottom row are, respectively, mean bias, root mean square error and Pearson correlation coefficient, computed using data between 75S and 75N.".

4. Figs. 3, 4, 5, 8, 9, 10, and 11: Summary statistics, i.e., mean bias, rmse, correlation coefficients, for the differences between model and observations would be helpful. On Figs. 6 and 7, the bias is evident, but rmse and correlation coefficients would provide valuable additional information about the fidelity of the model patterns to the CERES observations.

We thank the reviewer for this suggestion. We have adjusted the figures so that summary statistics are present, which shows for most quantities that CanAM5 performances as well as or better than CanAM4. We only added this information for plots for 2D horizontal field, e.g., top of atmosphere or surface, since the statistics can be difficult to interpret for latitude-pressure plots (Figures 8 and 9).

Information was added as follows,

- Figure 3: Mean bias, rmse and correlation coefficient have been added to the plots. These statistics are computed for data between 60S and 60N to be consistent with the range used for the zonal means.
- Figure 4: Mean bias, rmse and correlation coefficient have been added to the zonal plots. These statistics are computed for data between 75S and 75N to be consistent with the range used for the zonal means.
- Figure 5: Mean bias, rmse and correlation coefficient have been added to the difference plots.
- Figure 6: An additional row has been added to the figure which shows the rmse and correlation coefficients.
- Figure 7: Rather than adding the statistics to these plots, which in our opinion would make it very difficult to read, we have added a new figure based on Figure 7 to the supplemental that shows the mean bias, rmse and correlation coefficient (Fig. R1). The text has been adjusted to refer to this new figure.
- Figure 10: Mean bias, rmse and correlation coefficient have been added to the difference plots.
- Figure 11: Mean bias, rmse and correlation coefficient have been added to the difference plots.
- Figure A1: Mean bias, rmse and correlation coefficient have been added to the difference plots.

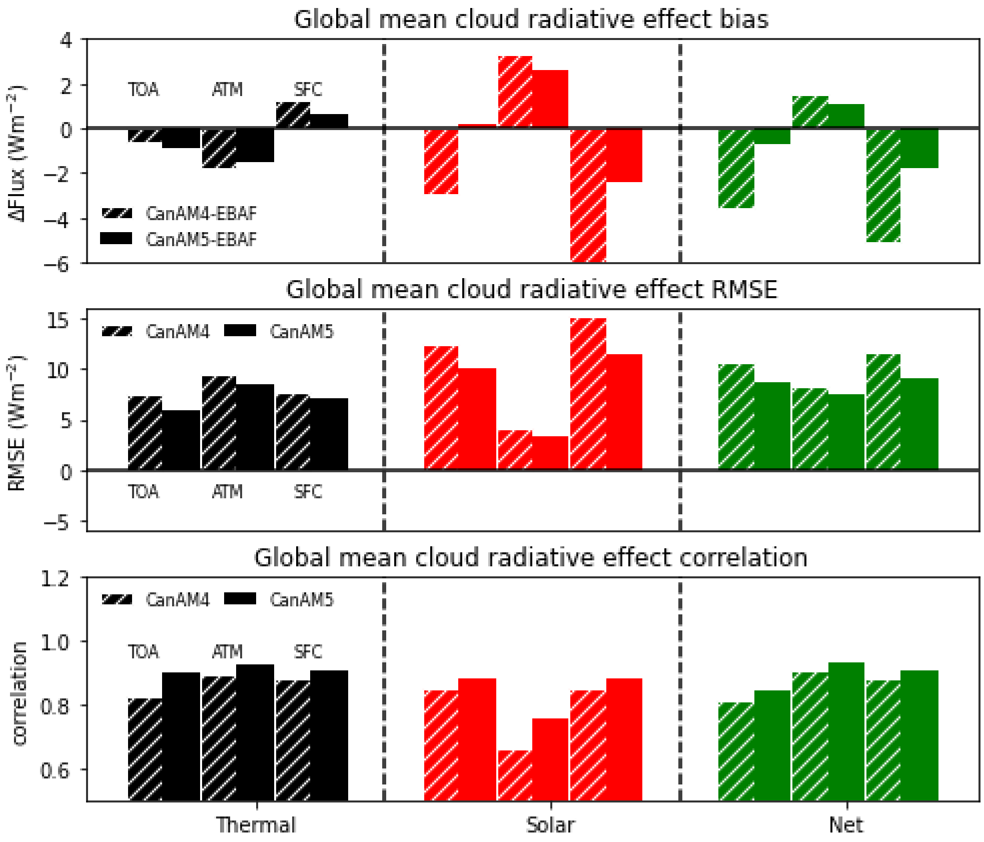

Figure R1. Global mean bias, RMSE and Pearson correlation coefficient for cloud radiative effects.

5. pp. 18-19, ll. 13-1, Fig. 9: Regarding zonal-mean temperature differences, note that north of 60N in DJF and MAM a large fraction of the space has oppositely signed differences.

True. We have added a sentence to the text noting this,

"However, there are regional and seasonal differences, for example, temperatures between December and May poleward of 60^{o}N in the stratosphere are not as systematically biased warm in CanAM5 relative to CanAM4."

6. p. 21, ll. 1-2: The text states that TOA net downward flux was tuned to produce a reasonable 1850 control for CanESM5. Provide a brief characterization of this simulation, i.e., how well is the TOA radiation balanced and what, if any, drifts are occurring?

We have added a reference to Swart et al., 2019 which shows these quantities in the CanESM5 preindustrial control simulation. As noted in the Swart et al., 2019 paper there are drifts, as are the case for most models participating in CMIP6, but the magnitude of the drifts is significantly smaller than changes over transient CMIP6 experiments.

7. pp. 14-16, ll. 9-4, and p. 21, ll. 1-4: The TOA net LW+SW imbalance (Earth Energy Imbalance, EEI) in CanAM5 of 3.1 W m_{-2} is quite large relative to the CERES EBAF value of 0.9 W m_{-2} and the IPCC-estimated total anthropogenic 1750-2011 radiative forcing of 2.3 W m_{-2} . This indicates significant errors in the *model's ability to simulate the observed energy imbalance of the Earth-atmosphere system given realistic boundary conditions. Retuning to produce a stable coupled integration is effectively a flux adjustment, even if not explicitly applied as such. Alternatively, these tunings can be viewed as introducing compensating errors in the coupled model to correct whatever deficiencies lead to the drift or unrealistic simulations there. The behavior of the coupled model using tunings which produce the observed EEI in an AMIP integration of CanAM5 would provide an informative gauge of the seriousness of these model deficiencies. I would encourage considering showing a measure of the problematic CanESM5 simulations using an uncoupled atmospheric configuration with a realistic EEI. The revised text should acknowledge the importance of these deficiencies.*

We have adjusted the text to make our point more clearly in Sections 4 and 6.2. As noted in the original text, when tuning we were targeting a stable climate that satisfies a handful of targets, e.g., the global mean temperature and sea-ice area. The properties of CanESM5 for a pre-industrial experiment is shown in Figure 1 of Swart et al, 2019.

As part of development, initial versions of the coupled model were tested with a variety of tunings that roughly matched the observed EEI in present-day AMIP simulations of CanAM5. However, pre-industrial simulations with CanESM5 using such tunings produced colder, unacceptable global mean near surface temperatures, by up to 2 K, and excessive sea-ice. While our attempts were not exhaustive, we came to the same conclusion as the reviewer that we were faced with a model structural error, and at the time we were curious what the EEI in present-day AMIP simulations was for other CMIP models.

The reviewer's comment has led us to reconsider the question of just how pervasive this type of structural error is in CMIP6 models. Now that CMIP6 is essentially concluded, we can analyze the output of CMIP6 models to investigate. In the Fig. R2 we have plotted the global mean net downward flux averaged over the period 2003-2009 for all CMIP6 models that have at least one "amip" and one "historical" simulation. Red dots indicate present-day amip EEI values while black dots indicate historical values. Interestingly, many CMIP6 models simulate a net flux at TOA that is well outside CERES observations in present-day AMIP configuration (some exceptionally so) but all models are generally less likely to be outside CERES when run in coupled configuration. From this figure it is clear that such model structural error is quite pervasive across CMIP6 models and that CanESM5 is not an unusual member of this group. As we are not aware of such analysis in the literature, we are motivated to publish these results in a follow-on study to highlight this issue. We thank the reviewer for reminding us of this question.

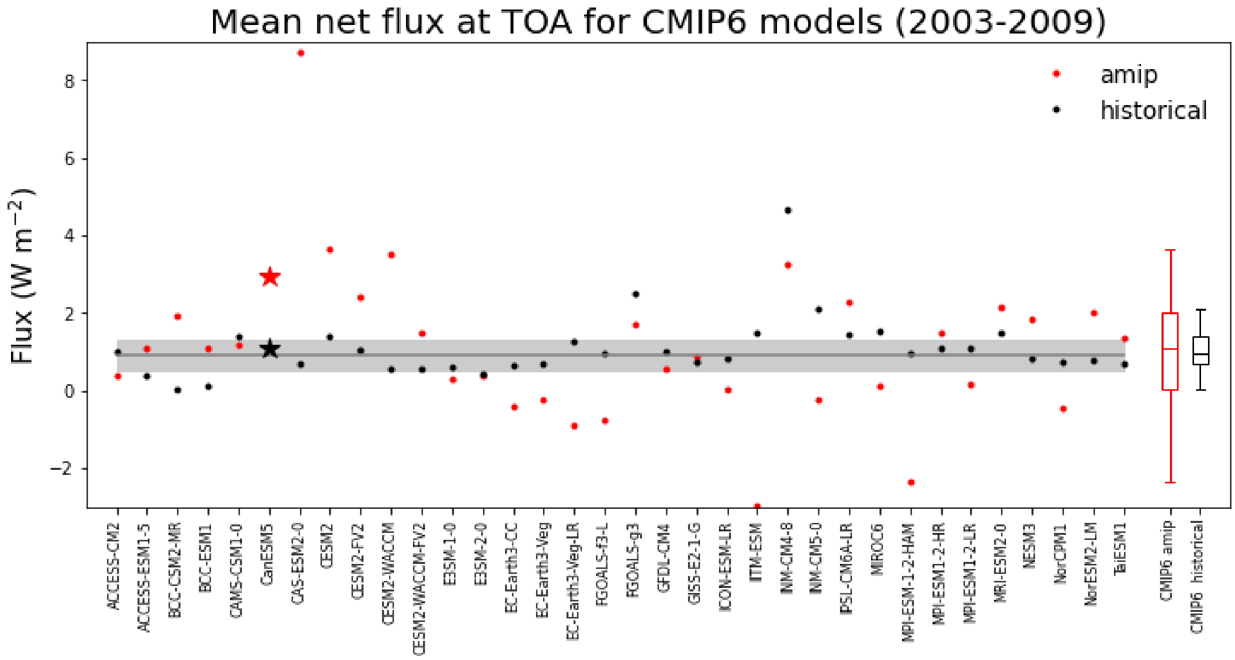

Figure R2. Net flux at TOA averaged over 2003-2009 for CMIP6 models using the "amip" experiment (red) and the coupled "historical" experiment (black).

Technical Corrections

p. 2, l. 17: “The the” -> “The”

p. 2, l. 18: “tropopause” -> “troposphere”

p. 4., l. 11: “CanAM5” repeated.

p. 16: l. 13: “shortwave the” -> “shortwave at the”

p. 19, l. 5: Fig. 11 shows CanAM5/4, not CanESM5/2.

All of these corrections have been made in the text.