Response to Reviewer I (Isaac Held) (Author Comments in Blue)

***Interactive comment on* “Radiative-Convective Equilibrium Model Intercomparison Project” *by* Allison A. Wing et al.**

**I. Held (Referee)**

isaac.held@noaa.gov  
Received and published: 24 September 2017

I am very supportive of the proposed RCE intercomparison project. The paper is generally well-written, and my comments primarily relate to the specifics of the proposed project.

We thank Dr. Held for his support of RCEMIP and constructive comments regarding the specifications of the proposed simulations. We have made several changes to the details of the simulation specifications in response (see revised manuscript and responses to individual comments below). In addition, we hope that the revised manuscript (see especially Section 6) better conveys our goal with regards to the simulation design: to keep it as simple as possible in order to maximize participation, establish a baseline, and inform subsequent experimentation. We envision that RCEMIP will eventually extend far beyond the simulations laid out in this paper, as we recognize that the baseline we propose will not be a definitive representation of the RCE climate for many of the reasons raised by the reviewers. While it is clear that certain physical parameterizations (or lack thereof) may lead to biases that require further investigation, we see the simulations proposed as a way to bring the community together to get us to that next point.

Major issues:

I continue to be puzzled that the papers on RCE simulations with CRMs hardly ever discuss the sensitivity of the results to boundary layer formulations. In fact, most of these simulations have no boundary layer closure at all except for the Monin-Obukhov (M-O) surface flux framework. M-O-is, in fact, a boundary layer formulation, and if that is all that one has in the model, one is (arbitrarily it seems to me) assuming that a closure is needed between the surface and the lowest model layer but not between the first model layer and the second model layer. The specification of a minimum wind speed in the bulk surface flux formula can also be thought of as a boundary layer turbulence parameterization of sorts, and the fact that you are careful to include this in your specification is implying that you don’t trust the resolved scales to provide the right gustiness. Running at 3km resolution without a boundary layer above the lowest model level strikes me as extreme (not that I haven’t done it myself). I would really like to see this topic directly discussed by the authors. This is a “grey zone” that seems to get less attention than the problem of transitioning from parameterized “convection” to explicit “convection” and one that I find very confusing.

We agree that the lack of boundary layer formulations in many CRMs could be a problem and has been underexplored. We anticipate that some of the CRMs participating in RCEMIP will employ boundary layer closures while others will not; it will be informative to see whether those without boundary layer closures consistently display different behavior from those that do. If so, targeted sensitivity experiments of boundary layer formulations could be a focus of phase two of RCEMIP. We note that our inclusion of a specification for a minimum wind speed in the surface flux calculation was focused on making sure that everyone uses the *same* minimum wind speed. We now mention that models should use their default boundary layer closure, if one is employed, in our specification of the surface boundary condition (Section 3.2.1), and mention boundary layer closures as a possible focus in phase two RCEMIP simulations (Section 6).

Along the same lines, it would be nice if the text more explicitly encouraged much higher resolution simulations than 1km in the 100km domain, closer to LES resolution for the planetary boundary layer.

We agree that it would be great if modeling groups were able to complete additional simulations with higher resolution. The simulations described in the paper are only the *required* simulations – we hope that participants will be eager to go beyond them, and now explicitly state that we encourage this (see Section 3.3, first paragraph). We would like everyone to complete simulations at the 1 km resolution in the 100 km domain so that they can be compared without worrying about the impact of varying resolution. We now also explicitly encourage LES simulations in Section 3.3.5.

I would strongly encourage you to reconsider and ask groups to run with a standard microphysical mechanism in addition to their model’s microphysics. Otherwise, there is a good chance that the diversity of simulations will be dominated by the diversity of microphysical assumptions. (For example, we know that in RCEs assumptions about ice fall-speeds will exert a strong control on the cirrus climate.) I realize that this would be work for the authors, who have to be confident that this standard microphysical mechanism worked reasonably well in a couple of the models that they are familiar with. The proposal here is to leave this issue to a second round, but I think this is postponing the inevitable. (For example, you suggest that models with more elaborate microphysics, with explicit activation of CCNs perhaps, be adjusted to create simulations that produce the same droplet number concentrations as recommended for models that specify these droplet numbers, but what if droplet numbers change significantly with warming in these models?)

We agree that large differences could result from differences in microphysics, and that it might be correct that the diversity of simulations will be dominated by the diversity of microphysical assumptions. However, we think that it is useful to first determine the full range of RCE simulations and *then* proceed to test the microphysics sensitivity by imposing a simple microphysics scheme on all models in the second phase of RCEMIP. In the past, groups have found large sensitivities to microphysics, but that might also reflect that the behavior of microphysical parameterizations are easiest to change (i.e., it is easy to modify a fall speed, but harder to change an underlying assumption in a boundary layer representation, for instance). In addition, it might not make sense to specify the microphysics without specifying the treatment of cloud optical properties (radiation), representation of partial cloudiness, etc…, and this would be too much to accomplish with our first set of simulations. Importantly, our goal with the first phase of RCEMIP is to keep the required simulations to a minimum and as close as possible to a models “standard” configuration so as to encourage maximum possible participation and limit the possibility of new physics introducing bugs that are not characteristic of a specific model. Perhaps it is “postponing the inevitable” but we believe that it is worth first providing a framework and taking stock of where things stand. We think that determining, for example, how many of the models have a decrease in high cloud fraction with warming with the “standard” configuration is valuable (if hard to disentangle), because presumably all the different schemes used are individually reasonable and justified choices and we don’t want to immediately bias the results in the direction of one scheme over another.

Regarding the specification of droplet number concentrations we suggested, we indicated that models should use the standard from their Aquaplanet configuration, and only provided numbers so that those without an Aquaplanet configuration would have some guidance as to what to choose.

The utility of single column models, in my view, is dependent on their “effective stability”: the single column equilibrium, whether time-independent or dependent, with no resolved flow is always a solution of a fully doubly periodic box or global RCE model that uses the same column physics (in fact, this is a good test that a global model has been successfully homogenized). But it is invariably spatially unstable. To the extent that the time average of the RCE state in the 100km box bears little resemblance to the single column equilibrium, the single column results are going to be hard to relate to the rest of the project. I would try to encourage any contributions of single column RCE simulations to also provide full doubly-periodic 100km box simulations for comparison, or at least global RCE simulations if these are easier to perform.

These are important issues to consider when interpreting the results of the simulations across different model types, but we do not believe precludes the inclusion of single column models in RCEMIP. In theory, the time average of a doubly periodic box or global simulation should be the same as that of single column models; the exception is if the box or global simulation aggregates (we assume this is the spatial instability referred to above). We acknowledge that, in general, it is difficult to compare results from models that aggregate and models that do not, but we anticipate that many of the 100 km simulations at CRM resolution will not aggregate, which then should be able to be compared to the single column simulations (disagreements would result from deficiencies in either or both models, which such an intercomparison might reveal). We anticipate that most single column models will be paired with a corresponding global RCE simulation (as this is the initialization procedure for GCMs), as was suggested. And, at minimum, single column models will be able to be compared to other single column models. Overall, we simply chose to design the experiments in a way that could support take-up of RCE simulations by the single column community, should they be interested in using our framework. To clarify the SCM set-up and its utility, we have added a section about single column models in Section 3.3.4.

For the same reason, I don’t particularly care for initializing global models with single column equilibria. Among other issues, what if these equilibria are time-dependent? (This is not hypothetical – it may be the rule rather than the exception with cloud- radiative interactions present). Do you initialize all the grid point with the same phase?

Part of the motivation for initializing global models with single column equilibria is to have a consistently configured “control” simulation that does not have convective aggregation to compare the global simulations to. All model grid points would be initialized with the same sounding from the single column model, but then small random noise will be added to break the symmetry. The alternative is to have global models initialize from an externally specified sounding, but we do not think this is better than using potentially time-dependent equilibria, plus, it would then be unclear how to create a control simulation without aggregation to compare to. Preliminary tests with CAM5 indicate that the RCE state achieved is not sensitive to how the model is initialized, leading us to believe that the potential issues of using single column equilibria to initialize global models will not be of first order importance.

I would not specify number of grid points to be used. I would just specify the physical size of the domain and then encourage simulations with 1km resolution in the small domain and 3km in the bowling alley domain, while also encouraging any other resolutions that the groups are interested in exploring.

We would like all models to use the same domain size and resolution so that we have a consistent configuration that can be compared across models. Given that we know that at least some aspects of RCE are sensitive to domain size and resolution (i.e., Muller and Held 2012), we do not want to conflate that sensitivity with model-model differences. We certainly do encourage additional simulations at other resolutions, with the same physical size of the domain. But we agree that it is probably better to specify the physical size rather than the number of grid points, so we have reworded the domain specification for CRMs (Section 3.2.1) to an approximate physical domain size and grid spacing (see lines 32-32, page 9 and lines 1-2, page 11).

The recommended upper boundary seems quite low to me, making it awkward to look at the vertically propagating gravity waves in the stratosphere excited by the convection. This is a nice opportunity to look at the robustness of the simulation of those waves, especially in the bowling alley case. On a related point, I would be a little worried that the bowling alley configuration will in some cases allow the generation of QBO-like oscillations, as in the two-dimensional limit. This could cause some large unexpected differences between simulations. It is worth mentioning at least.

We have changed the recommended CRM vertical grid to have a minimum of 74 levels (up to 33 km) (See line 26, page 10). The 4096 x 64 bowling alley configuration used in Wing and Cronin (2016) did exhibit QBO-like oscillations, as noted in that paper, although the effect on the troposphere was minimal. The 2048 x 128 simulation examined in that paper (which is the domain geometry to be used in RCEMIP) did not experience such strong oscillations.

You seem to be initializing stratospheric water with extremely small, but non-zero, mixing ratios. But won’t stratospheric water take forever to equilibrate in RCE simulations, where the only thing that can mix vapor up is breaking gravity waves? This vapor has some effect on sensitivity of radiative fluxes to warming. Kuang and Bretherton, JAS, 2004 have circumvented this problem by specifying a small vertical motion (a Brewer-Dobson upwelling extending through the troposphere and stratosphere), which allows the tropopause cold-trap to operate. As in nature this also makes it easier to study the tropical tropopause layer, where convection and upwelling can both be important. In any case, this issue of equilibrating stratospheric water should be discussed. (How do you get water vapor into the stratosphere in a single column model? Do you intend that these models include sub-grid closures for mixing by convectively generated waves?)

We examined the evolution of stratosphere water vapor in the preliminary CAM5 simulations and SAM simulations at several different model levels. In the CAM5 simulations, stratospheric water vapor is indeed not equilibriated by the end of year 3, but we believe this has little impact on the RCE state in the troposphere; the CAM5 simulations are in equilibrium in terms of precipitation, outgoing longwave radiation, and omega. In the cloud-permitting SAM simulations, the water vapor in the lower stratosphere seems to be approximately equilibrated within 100 days. Regarding the motivation for initializing the stratospheric water with very small but non-zero mixing ratios, we followed the procedure of Reed and Jablonowski (2011). We do not think that there is an obvious way to deal with the stratospheric water vapor correctly, and so just acknowledge in the paper that the lack of equilibration will be an issue and will be monitored and assessed in the evaluation of the simulations (see lines 15-18 on page 9).

Minor comments:

p-3, l-2. Held, Zhou, Wyman, 2007 also compared the cloud feedbacks in RCE against those in the tropics of a global model with the same column physics.

We now reference this paper (see line 8, page 3) and apologize for the omission.

p.-5, l-20. Sorry, but false precision is a pet peeve of mine. You specify the mean surface pressure to 6 significant figures, but the mass of the atmosphere is not conserved, it increase with warming due to increase in vapor. Some global models correctly incorporate the mass source/sink of vapor, and these models will just ignore the value that you specify, they need to specify the dry mass.

This has been removed in the revised manuscript.

As another example, in the list of constants you are careful to specify the value of the latent heat of vaporization, but you don’t specify its temperature dependence – or, equivalently, the heat capacity of the condensate. Are you expecting the latter to be set to zero? You specify the (mean) radius of the Earth with precision and then suggest that models just use smaller radii, a good idea since the radius in the non-rotating case is totally irrelevant until it approaches the depth of the model domain. Meanwhile differences in assumptions about surface smoothness or ice fall-velocities or many other things that are literally orders of magnitude more important are passed over. The table of geophysical constants is pointless in my opinion; just ask modelers to use their standard Earth values.

Our motivation for including the table of geophysical constants was to follow the convention of other MIPs (i.e., the Aquaplanet Experiment) in providing a table of these values. We are, in fact, asking modelers to use the standard Earth values for their Aquaplanet configuration. We specified the value of the latent heat of vaporization at 0 C; models should follow their usual formulation for its temperature dependence.

I would not ask the modeling group to provide your preferred measures of aggregation. Other measures might turn out to be more interesting. These should be relatively easy to generate in a uniform way after the fact from the data archive. You should keep the post-processing requested from the groups to a minimum. (Some groups may not be interested in the dynamics of aggregation.)

We agree that the post-processing requests should be kept to a minimum and the only diagnostics that must be computed online are cloud fraction and the moist static energy budget terms. We have removed the mention of the autocorrelation length, but retain a shortened description of the two aggregation metrics, since we show plots of one of these quantities in the paper. We have rephrased the text to indicate that code to compute those metrics will be available on the website.

As for the energy budget, I am not sure why you want this in advective rather than flux form (I may be missing something.). A data request could include the fluxes of energy at the boundaries of the grid cells – at least in CRMs and grid-point GCMs. These are more likely to be precisely defined by the underlying numerical scheme than the advective form.

Previous studies have also calculated the moist static energy budget in models using the advective form (e.g., Anderson and Kuang 2012, J. Climate). To rewrite it in flux form, one needs to assume something about the continuity equation, which takes different forms in different models depending on which version of the equations of motion they use. In the end, the estimation of the advection term is going to be inherently uncertain and we just ask groups to do the best they can.

I can’t see anything in Fig. 6

We are not sure what is meant by this comment; when we download the PDF of the paper from the GMD website, Figure 6 is visible. It shows the daily average column water vapor at the beginning and end of the small domain SAM simulation, which is horizontally homogeneous because the simulations are unaggregated and the convection is quasi-randomly distributed, and thus “averages out”.