-------------------

Reviewer #3: I thought this was very nice and had an interesting analysis of the differences. However,I have a philosophical problem with the description "effect of different dynamical cores" which occurs in the title and throughout the paper. To me this implies that the fundamental cause of the differences is the dynamical core and that the two cores are not capable of producing the same simulation. That is not clearly shown in the paper and I believe that that may not be the case. I think the cause is more a difference in "effective resolution" associated with the damping characteristics of the numerical schemes and that there might be resolutions that yield the same simulation statistics, at least for the large scales that are considered here. It is conceivable that schemes might not produce the same statistics of the large scale and at the same time of the scales near the truncation limit, but those smaller scales are not considered here. Currently I think the term "effect of different dynamical cores" is misleading and should not be used.

We know that if we take two different cores and run them at vastly different resolutions we will get different climate. Also if we take one core and run it at different resolutions we will also get different climates. To me these differences are not attributable to the dynamic cores but to the resolution.

The paper makes only a token attempt at looking at resolution. In the primary comparison in the paper between IAP and CAM, both are applied on a 256 (lon) by 128 (lat) point grid. Given the vastly different accuracies of the approximations in the two models we would not expect these cores to produce similar results when applied to the same grid. The authors also show a result of IAP run on a 1 degree (360 by 180 point) grid. These resolutions are not very different. However, the higher resolution IAP climate has moved closer to that of CAM. Williamson (2008) showed that T85 spectral aqua-planet climate was comparable to that of the finite volume at 1 degree as the authors state. But their finite difference scheme is not as accurate as the finite volume, so they would likely need finer resolution still. They should run a simulation of at least 0.5 degree to see if the IAP climate is closer to CAM's.

Plots of the kinetic energy spectra should be included to provide some indication of the damping characteristics of the IAP at the different resolutions. It would be nice to include CAM on the same plot.

Perhaps there is a more serious problem, related to the CAM time step. Page 10, line 16 says the time step used for the dynamical core and physics is 600s. 600s is the standard time step (DT) for T85 CAM. But CAM uses three-time-level centered differences and the physics is calculated over 2DT or 1200s. Page 10, line 4 says that the IAP core uses a 600s time step for the dynamical core and physics. The IAP core seems to use a two-time-level scheme in which case the physics might very well be calculated with a 600s time step. Thus the physics might not be calculated with the same time step with the two cores. The CAM physics is very sensitive to time step and this would not be a proper comparison. If this is the case the IAP will need to be redone with a 1200s time step. The CAM should not be run with a different time step than the standard because it is being compared here with the atmosphere, and the standard, tuned-released version only should be used for such comparisons.

Minor points:

Page 8, line 15-16: Fourier filtering is used to damp the short longitudinal wavelengths, which happen to be high temporal frequency might be a better way to express it.

Page 10, line 20: the spectral transform is used for the non-linear terms, not just quadratic terms.

Page 10, line 20: What does "most of the interior of the atmosphere" mean?

page 11, line 18: "performance is similar to that of CAM3.1" seems rather vague. What is meant here, especially as the paper concentrates on some seemingly fundamental differences,

page 12, line 8: "colder ... by 1-2 K". I do not see this, since the figure only has a 1K contour, but not a 1.5 K contour.

Page 13, 1st paragraph. The transient eddy statistics are calculated differently for the model and for the reanalysis. I would like the potential differences to be quantified to clearly establish that the difference is not important. (Or reference such a calculation already shown elsewhere.) This can easily be done by applying the two methods to the model data.