Manuscript # JOC-14-0316
Reviewer: 1

Comments to the Author
This paper, which has its genesis in a similar exposition by Frank (2008) in 'Skeptic' magazine and has gone through a number of previous submissions to the technical literature. I have previously reviewed this manuscript and since the submission has not changed in any important way, my review here is necessarily similar to my earlier one. I note a few of the more trivial errors have been corrected, but the fundamental conceptual problems remain.

The author purports to emulate the result of GCM calculations under specific scenarios with a linear, instantaneous function of the forcings. This is only successful in situations where the forcing is basically linear since the thermal inertia of the climate system is ignored. From an emulation of the GCM temperature projections, the author derives a completely erroneous uncertainty due to cloud forcing biases and then purports to 'propagate' errors from the mean state energy flux into absolutely nonsensical (and self-evidently wrong) temperature projections of the future and past. This is fundamentally flawed and cannot be remedied to make it publishable.

Overall comments:

The confusion between errors in the mean state and errors associated with a perturbation are pervasive (equivalent to conflating the error in the constant and first derivative of a complex function). See below.

The author's claim that published projections do not include 'propagated errors' is fundamentally flawed. It is clearly the case that the model ensemble may have structural errors that bias the projections, but these are not derivable in the manner the author claims. Indeed, he has not demonstrated that any of the errors in the climatology that he highlights are even correlated to differing model outcomes, let alone dominant sources of projection spread.

The application of naive error propagation theory for a compound measurement to this case completely ignores the feedbacks and complexity of the models that render any assumption of statistical independence between subsequent errors null and void.

The demonstration of the fallaciousness of the reasoning demonstrated here is obvious upon assessing any other climate change simulated by the models. For instance, looking the last glacial maximum, the same models produce global mean changes of between 4 and 6 degrees colder than the pre-industrial. If the conclusions of this paper were correct, this spread (being so much smaller than the estimated errors of +/- 15 deg C) would be nothing short of miraculous. It isn't because they aren't.

Going more deeply into the methodology here, equation 6 (p 18) is the key assumption, that temperature changes are the instantaneous response to the normalised forcing relative to some t0. However, by the logic so far, it is clear that this delta(T) does not include any feedbacks to radiative forcing (since both the '0.42'was calculated holding everything else constant). The F0 is in the right ballpark for the radiative forcing associated with complete removal of the main GHGs. The formula give 13.9 K for the pre-industrial climate (i.e. 1900 in this case), and so the difference in temperature from the pre-industrial is given by Tanom=13.9 * Fanom/F0 = 0.41 * Fanom. Note that this gives a climate sensitivity to 2xCO2 (Fanom=4W/m2) of ~1.6C, somewhat higher than the expected 1.2 deg C no-feedback response generally estimated. The match of this PWM to the GCM output is simply because the transient climate response (TCR) (which takes into account ocean thermal inertia etc.) is close to 1.6 C, but this is entirely fortuitous.

Once the author decides to fit his model to the individual models, he is simply calculating an empirical TCR and all of the foregoing justification is moot. Of course, this will only work with scenarios that have roughly linearly increasing forcings. Any stabilisation or addition of large transients (such as volcanoes) will cause the mismatch between this emulator and the underlying GCM to be obvious.

The author than takes the uncertainty in the TOA energy fluxes (for which he uses +/- 4W/m2) and then assumes that this is also the uncertainty in the annual forcings. This is clearly nonsense, precisely for the reasons he earlier mentioned - i.e the cloud forcing errors are systematic - not random. Assuming that there is an additional random component in the forcings of this size produces error bars that are effectively a random walk and therefore will increase without bound over time. This neither matches what the models actually do, nor is it physically justifiable. (For instance, even after forcings have stabilised, this analysis would predict that the models will swing ever more wildly between snowball and runaway greenhouse states. Which, it should be obvious, does not actually happen). Given that the PWM is supposed to be an emulator of the GCM results, this a-physical result somewhat undermines its utility.

I will give (again) one simple example of why this whole exercise is a waste of time. Take a simple energy balance model, solar in, long wave out, single layer atmosphere, albedo and greenhouse effect. i.e. sigma Ts^4 = S (1-a) /(1 -lambda/2) where lambda is the atmospheric emissivity, a is the albedo (0.7), S the incident solar flux (340 W/m2), sigma is the SB coefficient and Ts is the surface temperature (288K). The sensitivity of this model to an increase in lambda of 0.02 (which gives a 4 W/m2 forcing) is 1.19 deg C (assuming no feedbacks on lambda or a). The sensitivity of an erroneous model with an error in the albedo of 0.012 (which gives a 4 W/m2 SW TOA flux error) to exactly the same forcing is 1.18 deg C. This the difference that a systematic bias makes to the sensitivity is two orders of magnitude less than the effect of the perturbation. The author's equating of the response error to the bias error even in such a simple model is orders of magnitude wrong. It is exactly the same with his GCM emulator.

The summary conclusion section is based wholly on the unsupported and erroneous results from the previous sections and does not need to be reviewed in depth.

Minor comments:

p3. l18. IPCC AR4 high end temperatures are twice this.

p6. l37. The author is simply asserting that uncertainties in published estimates are not 'physically valid' - an opinion that is not widely shared.

p6. l44. This is not true. Structural uncertainties (which go beyond individual parameter uncertainties) are very often assessed, most recently in the AR5 report. 'Systematic energy flux errors' are not a specific thing which can be propagated either.

p9. l48. This is fundamentally wrong. Calculating the magnitude of the greenhouse effect by removing a single absorber (in the case CO_2) does not produce the effect of what remains due to the spectral overlaps. The paper by Schmidt et al (JGR, 2010) shows clearly that this will underestimate the effect of CO_2 (i.e. their Table 1).

p10 to p13: this is all completely irrelevant for what is actually being done. Firstly, the whole calculation is contradictory to the earlier claim that this paper is purely concerned with GCM results and not the real world. Secondly, the answer would have been better derived directly from sensitivity studies from radiative/convective models themselves (ie. Ramanathan and Coakley, 1993). Third, the answer is wrong because of the neglect of spectral overlaps (which are as present in the models as they are in the real world).

p16. l47. There is no 'asymptote' at 1ppm CO_2.

p17. l17-40. All of the calculation here is just wrong. First, the attempt is being made to calculate the no-feedback contributions to the 33K GHE. The results of Hansen et al (1988) are applicable here and show that about 7K of the 33K is directly attributable to the single factor removal of CO_2. Allowing for overlaps (as discussed in Schmidt et al (2010)), you would get a slightly larger value of ~10K. Note that percentage attribution of the temperature change is not the same as for the LW absorption (because of the non-linear dependence of LW on T^4).

p18 l51. The author's definition of the 'water-vapour enhanced' CO_2 forcing confuses a forcing with the feedback and is fundamentally confused. Since it is a forcing, the temperature is fixed for this calculation, therefore so is the relative humidity and therefore the answer is exactly the same as for the CO_2 forcing itself. The description of the results from Lacis et al (2010) is wrong. Their result is that ~20% of the GHE is attributable to CO_2, and this has nothing to do with the model's climate sensitivity.

p21. l20-41: this is complete nonsense. The author is equating the attribution of the climatological GHE to the changes of GHE as a function of the feedbacks. This makes no sense whatsoever.

p24 l34: that models have systematic biases is well known (marine stratus decks, double ITCZs, excessive southern ocean absorption). I have no idea what 'theory-bias' has to add to the already extensive literature on why these biases exist.

p 26. Fig 4: this is completely irrelevant. Of course a smooth function over latitude with perhaps 2 or 3 degrees of freedom will show latitudinal correlations.

p27 l6: why not just look at the spatial errors? of course they are non-random!

p36 l27-37: this is the main paper error complete in one line. i.e. errors in the mean are not the same as errors in the perturbations.

p37 l20-25: Of course the AR5 mentions that it is plotting anomalies and what the baselines are.

p37. l39 onwards: I fail to see how 'theory-bias' as a concept leads to a specific conclusion that models are partitioning energy incorrectly among climate modes.

p40 l46 etc. All of this assumes that the errors always add and never interact. This is not valid for GCMs given the large negative feedback associated with the Planck response. This reduces the magnitude of error growth significantly.

p42-3 etc. This is patently absurd. The scenarios A, B and C are clearly distinguishable and make clear predictions as a function of the different scenarios.

p45. This response to earlier criticisms is very poor. The author claims that models are unable to predict GMSAT better than +/- 15 C because of propagation of errors associated with clouds. Yet models do not oscillate with variances that large. The author now suggests that these errors are not random in time, but rather systematic (constant in time), and yet, the estimate of LGM temperature differences (as assessed above) are all within a couple of degrees, an order of magnitude smaller than the author's claim. Therefore if the error is neither random in time nor systematic, where pray is it to be found?