To: pfrank830@earthlink.net

#### 02-Sep-2014

### Dear Dr Frank

Manuscript # JOC-14-0316 entitled "Propagation of Error and the Reliability of Global Air Temperature Projections" which you submitted to the International Journal of Climatology, has been reviewed. The comments of the referees, all of whom are leading international experts in this field, are included with this letter.

As you can see, two of three reviewers found fundamental flaws in your manuscript. It has, therefore, been denied publication in the International Journal of Climatology.

Sincerely,

Dr Radan Huth International Journal of Climatology huth@ufa.cas.cz

Referee(s)' Comments to Author:

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. I18. IPCC AR4 high end temperatures are twice this.

p6. I37. The author is simply asserting that uncertainties in published estimates are not 'physically valid' - an opinion that is not widely shared.

p6. I44. 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. I48. This is fundamentally wrong. Calculating the magnitude of the greenhouse effect by removing a single absorber (in the case CO2) 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 CO2 (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. I47. There is no 'asymptote' at 1ppm CO2.

p17. I17-40. 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 CO2. 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 I51. The author's definition of the 'water-vapour enhanced' CO2 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 CO2 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 CO2, and this has nothing to do with the model's climate sensitivity.

p21. I20-41: 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 I34: 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 I6: why not just look at the spatial errors? of course they are non-random!

p36 I27-37: 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. I39 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?

# Reviewer: 2

# Comments to the Author

I think that it is very interesting paper. If the claim is proven to be true, this paper will largely resound in the science community which depends on the reliability of long-term climate prediction. The claim that errors due to cloud bias are propagating and thus should be quantified based on accuracy (not precision) sounds perfectly like what is required in normal science. Thus I think that this paper needs to be published in the end.

There are some important questions to reassure that the claim of this paper is correct. The forcing error due to cloud bias may be damped by Stefan-Boltzmann feedback embedded in current climate models. The perturbed temperatures at time i due to cloud bias will be partly restored at time i+1 by the release or trap of energy by Stefan-Boltzmann feedback process. In addition, cloud fraction bias is not all, leading to error of ±4 W/m<sup>2</sup> of cloud forcing. Models have many different substances such as sea ice/snow, vegetation, cloud properties, and precipitation, etc. all of which also act to add error, or compensate error. Then even so, why are model-projected temperatures not too variant in the year of 2100 in Figure 5. I hope that the author can properly reflect my concerns in the manuscript, so readers can be more confident with the unique claim of this study.

Minor comments:

P2, L13: Water-vapor-enhanced CO2 forcing is not familiar to the general readers. Please explain it.

P2, L15: Unclear terms: error averages? Its high correlation among GCMs?

P3 to the end: Please follow the correct citation format for IJC throughout the manuscript. "Celsius. [IPCC, 2013; IPCC\_2007 et al.,

2007b]" should be corrected to "Celsius (IPCC, 2013; IPCC, 2007).".

P5, L47: Remove "(sure knowledge)".

P7, L5: "error thought model..."?

P7, L14: "water-vapor-enhanced (wve) greenhouse effect" was not explained, but explained later on P9, L17. Explain earlier when this term appears firstly.

P7, L32: "GHG" was not explained.

P9. L36: Remove "or absolute" since it is confused with "relative humidity".

P10 to the end: 15  $\mu$  should be replaced by 15  $\mu$ m.

P10 to the end: The first sentence should be placed in the next line of the subtitle.

P11, L56: Why "θ=45"?

P12, L57: "under ( ), clear sky, and; ( ), cloud cover" includes strange symbols.

P17, L56: "Model E"?

P19, L13: It was not clearly explained how  $\Delta$ Fi is calculated.

P19, L20: "ΔΔTa"?

P20, L45: "Figure 3:10 in [AchutaRao, K et al., 2004]"; please correct the format.

P25, L34: "The mean of the target MODIS and ISCCP2 A-train satellite observations"; how did you calculate the mean? For example, the data period and the method.

P25, L36: If possible, it is advisable not to use "Auxiliary Information", but to include all information within this manuscript.

P25, L55: I cannot understand what is "lag-1 latitudinal autocorrelation".

P30, L30: "up-radiant"? Do you mean "upward"?

P30. L49: I wonder why the authors did not consider SCF. SCF and LCF are compensating each other.

P30, L56: "is not separate" should be "is not separated".

P33, L51: "root-sum-square"; I wonder why not "root-mean-square" that is more popularly used.

P33: Please check equation 8; e i(T) seems constant.

P34: "descends directly from eqn. 2, eqns. 6&7, and Sections 2.2-2.4 above" is not correct format for IJC.

P34, L41&L54: "4AR" should be "AR4".

P44, L35: "Summary Conclusions" should be "Summary and Discussions".

P45, first paragraph: nobody knows previous version of this paper. I don't think this paragraph is necessary.

P46, L11: "evaluate" should be "to evaluate".

P47, L6: What is "Equation (10)"?

P49, L41: "poor resolution" sounds like spatial or time resolution. If it is not that meaning, please rephrase it.

P52. L44-48: 'degree' symbol is missing.

P53, L53: "(CMIP5) an AGW signal" seems like typo.

Reviewer: 3

Comments to the Author

If this paper is correct, it is a profoundly important paper. If this paper is flawed then to publish it would be profoundly misleading. The question is, which is it.

This paper shows that models cannot predict the weather years ahead of time, due to the accumulation of errors. This is already

known, so the result is not novel. We have known this formally since Lorenz. Unfortunately the author misunderstands this result and claims that the models cannot predict climate. The equation that he has applied, eq. (2), does not however apply to climate which is the average state of the atmosphere not the instantaneous state. It might conceivably apply to weather – and he proves weather is not predictable on longer timescales but it is a mis-application of Equation 2 to apply it to climate.

I think it is actually very arguable whether Equation (2) does apply to the weather state. The equation assumes that successive errors are independent and additive, yielding a random walk. This is not how dynamical systems behave, because in the short term errors grow exponentially (in effect, are multiplicative) and must be characterised not by a linear growth parameter as assumed in this manuscript, but by an exponent (the Lyapunov exponent). Thus the situation is even worse than the author believes for weather, and instead of months of predictability, as implied by Fig. 5 with linear error growth, there are only a couple of weeks of actual predictability. Over the long term, the dissipative character of dynamical systems (including the climate system) causes the errors to be bounded by the dimensions of the attractor. The author is strongly urged to consult a textbook on dynamical systems theory since the application of (2) is a very basic error. If he does not want to read dynamical text books by climate scientist, read some by applied mathematicians.

The author's calculation is completely inapplicable to climate modeling. Indeed if we carry such error propagation out for millennia we find that the uncertainty will eventually be larger than the absolute temperature of the Earth, a clear absurdity. In reality climate models have been tested on multicentennial time scales against paleoclimate data (see the most recent PMIP intercomparisons) and do reasonably well at simulating small Holocene climate variations, and even glacial-interglacial transitions. This is completely incompatible with the claimed results.

In summary, the claims made in this paper are profoundly important if they are right. Unfortunately, they are wrong at several important levels and the paper has to be outright rejected.