Patrick Frank 8 October 2017 Earth and Space Science Manuscript 2017EA000308 Response to Round 2 Review #3 (round 1 #6)

Summary Response.

This reviewer:

- 1. Shows no awareness that linear extrapolation of forcing is subject to linear propagation of error. This is the analytical core of the analysis and the reviewer never grasped it.
- Ignored the objective falsification of his claims in round one concerning [*John and Soden*, 2007] and [*Dessler*, 2013], items 1.2.1, 1.2.2, 3.2, 4.1-4.5, 5.1, 5.2.1, 6.1.2, 6.1.3, 6.2.1, and 7.6.2.
- 3. Contradicted his own round 1 case for rejection, item 7.1
- 4. Is wrong that, "[the long wave cloud] feedback differences ([in [Dessler, 2013])" are comparatively negligible, item 6.2.3.
- 5. Is wrong that model studies have validated equilibrium climate sensitivity, item 7.2.1
- 6. Is wrong that simulated climate change is free from uncertainty, items 2.2.1, 2.2.3, and 2.2.5
- 7. Wrongly presumed that similarity of model results is equivalent to no error, items 5.2.2, 6.1.3, and 6.1.4
- 8. Wrongly assumed that taking anomalies discharges uncertainty, items 2.2.1, 2.2.3-2.2.5, 3.1, 6.1.3, and 8.4.1.
- 9. Confuses physical error and statistical uncertainty, items 2.1.2, 3.3, 6.2.2.1, 8.3.2.1, and 8.4.1
- 10. Misconstrued the author's analysis so extraordinarily badly as to imply unprofessional negligence, items 2.2.1, 3.1, 7.6.1, 7.6.2, 7.6.3.1, 7.6.5, 7.6.6, and 8.1. See especially 7.6.1, 7.6.2, 7.6.5, and 7.6.6.
- 11. Does not understand calibration, items 2.2.4 and 7.5.2
- 12. Engages an incoherent logic, items 5.2.2, 6.1.4, 7.2.2, and 7.3
- 13. Confuses precision and accuracy, items 1.2.1, 4.1-4.5, 6.1.2, 6.1.4, 6.2.2.2, 6.2.3, 7.5.1, 7.5.2, 8.4.1
- 14. Does not understand error analysis, items 2.2.5, 3.1, 3.2, 4.5, 5.1, 6.1.3, 6.1.4, 8.3.1 8.3.2.1.
- 15. Derided what he clearly did not understand, item 8.5.1
- 16. Assumes a theory not in evidence, items 2.2.2, 2.2.6, and 7.2.3
- 17. Does not understand that fortuitous error cancellation does not remove uncertainty, items 2.2.4, 5.2.2, 6.1.3, and 8.3.2.1.

Extraordinary carelessness and severe analytical mistakes remove all critical merit from the review. It should be set aside.

The reviewer is quoted in italics below, followed by the indented response.

1.1 In my first review, I suggested that there were fundamental problems with the paper. Unfortunately, the author made no changes to the paper to address these concerns. He did write a response to my review, but I found his arguments to be extremely unconvincing. 1.1.1 The reviewer is thanked for being candid.

The reviewer evidently finds unconvincing that:

- 1. precision is not accuracy,
- 2. that model results ([*John and Soden*, 2007] Figure 2) can be (and were) falsified by comparison with observations,
- 3. that models cannot be used to validate models (i.e., as in [Dessler, 2013]) and
- 4. that observational standards are so poor that the model results ([*John and Soden*, 2007] Figure 2) were effectively meaningless.

As listed above and shown both below and in the round 1 response, the so-called "*fundamental problems*" are (and were) in the review rather than in the manuscript. There is no point to revising a manuscript in light of ill-conceived criticisms.

1.1.2 As shown thoroughly below, the reviewer found the authors arguments "*extremely unconvincing*" because the reviewer did not understand any of them.

It shown below that this reviewer does not understand, among other things:

- calibration,
- error propagation,
- the distinction between accuracy and precision,
- the difference between physical error and statistical uncertainty.

The fact that the author's arguments, not to mention the manuscript analysis itself, presumes a knowledge of these concepts that this reviewer clearly lacks, it should be no surprise that he found the author's arguments "*extremely unconvincing*." No one is convinced by arguments he does not understand, no matter how compelling.

- 1.2 Given this experience, I see little chance that the author is willing to make even the most cursory effort to address serious defects in the paper and I therefore reluctantly recommend REJECTION.
 - 1.2.1 The "*serious defects*" were in the review. This reviewer incorrectly asserted that [*John and Soden*, 2007] proved that models accurately simulated changes in climate.

The author showed that [*John and Soden*, 2007] concerned only model precision, and revealed nothing of simulation accuracy. This demonstration analytically disproved the reviewer's first rationale for rejection.

1.2.2 The reviewer then incorrectly asserted that [*Dessler*, 2013] showed good agreement between simulated and observed cloud feedbacks.

The author definitively showed that [*Dessler*, 2013] did no such thing, and had in fact deployed a circular argument to come to its conclusions.

The reviewer has chosen to ignore these refutations. This indicates little chance that the reviewer is willing to make even the most cursory effort to address the obvious deficiencies in his critical reasoning.

- 1.3 The serious defects were in the review, not in the manuscript. The author round 1 response demonstrated these defects. The reviewer has ignored them. Further evidence below indicates the reviewer did not even understand them.
- 2.1 Just to make sure we're on the same page, let me restate my primary problem with this paper: the crux of the paper is the key assumption stated on lines 590-594: "The rationale for eqn. 8 is straightforward. The response of the climate to increased CO_2 forcing includes the response of global cloud cover. However, global cloud cover is not simulated to better than average ±12.1 %. The resulting LWCF error means the magnitude of the change in tropospheric thermal energy flux in response to GHG forcing is not simulated to better than ±4 Wm⁻²."
 - 2.1 The reviewer's primary problem is evidently with [*Jiang et al.*, 2012] (±12.1% uncertainty in simulated global cloud cover) and with [*Lauer and Hamilton*, 2013] (±4 Wm⁻² uncertainty in simulated LWCF), rather than with the author's manuscript.

The reviewer clearly does not grasp that the quoted statement alone is sufficient to convey to any well-trained physical scientist that climate model simulations only poorly resolve the tropospheric thermal energy flux.

The inevitable conclusion must be that climate models are unable to resolve the thermal impact of CO_2 forcing.

2.2 What he's saying is this. Imagine you have a model that simulates some cloud coverage at some temperature C[T]. But climate change is a problem of derivatives, i.e., dC[T]/dT. The author assumes that the error in C[T] is equal to the error in dC[T]/dT. This assumption is key - without it, the entire paper collapses. The author makes no effort to prove it, but rather simply asserts it is true.

2.2.1 The author makes no such assumption. The author merely applies standard physical error analysis.

The reviewer wrongly assumes that calibration uncertainty is removed by differencing. If model calibration uncertainty is ±u, such that the simulations yield C₁±u, C₂±u, and if dC = (C₂-C₁), then the uncertainty in dC is $\pm \sigma_{dC} = \sqrt{u^2 + u^2} = \pm 1.4u$.

That is, the uncertainty in $dC_{2,1}$ is necessarily greater than the uncertainty in C_1 and C_2 . This is standard uncertainty analysis. The reviewer cannot escape it.

Review item 2.2 merely shows this reviewer does not understand uncertainty or its analysis, or the impact of uncertainty on calculated results.

2.2.2 Further, in asserting error-free anomalies, the reviewer is implicitly assuming that both the climate and the models follow linear response theory and that all simulation error is merely a constant single-sign offset. No evidence supports these assumptions.

The author addressed this in round 1 response item 1.4. However, the reviewer shows no

evidence of having read the author's response on this matter.

2.2.3 The author's position is that climate is simulated as state magnitudes C, T, not as slopes or anomalies, dC, dT (*cf.* manuscript lines 647ff). Error analysis follows from what is simulated, not from later inferences.

The uncertainties in simulation anomalies dC, dT are determined by the model calibration errors in simulated C, T. Note the distinction: C and T are the magnitudes that are actually simulated; dC and dT are the inferential simulation anomalies. The changes are calculated from the simulation; they are not directly simulated.

Calibration uncertainty propagates from the simulation, not from the changes, and \pm uncertainty does not subtract away. This last should be obvious merely by inspection of the arithmetic, i.e., $-\pm = \mp$.

The reviewer's $dC = C_2-C_1$. The simulation errors are unknown because future climate has no observables against which physical simulation error can be determined. The simulated states can therefore be conditioned only by the model calibration uncertainties.

Thus, simulations yield $C_1 \pm u_1$ and $C_2 \pm u_2$, where $\pm u$ is the calibration uncertainty in the simulated magnitude of C. Then dC = $(C_2 \pm u_2) - (C_1 \pm u_1)$ and the uncertainty $\pm \sigma_{dC} = \sqrt{u_1^2 + u_2^2} = \pm \sigma_{C_1, C_2}$, where subscripted $C_{1\nu}C_2$ on \pm sigma reference the uncertainty in dC. This is straightforward uncertainty analysis. The reviewer cannot escape it; nor can climate projections.

2.2.4 The reviewer is not correct in supposing the author assumes that $\varepsilon_{C[T]} = \varepsilon_{(dC[T]/dt)}$, where ε = error. The author's argument is not about physical error, ε . It is about uncertainty, $\pm u$. This difference is critical, and is apparently lost on the reviewer.

The reviewer's evident inability to recognize the difference between error and uncertainty is fatal to the review itself.

The $\pm 4 \text{ Wm}^{-2}$ is not physical error. It is a statistical uncertainty from model calibration. The $\pm 4 \text{ Wm}^{-2}$ does not subtract away. Nor is uncertainty removed by taking derivatives. Nor is uncertainty removed by fortuitous error-cancellation.

2.2.5 The reviewer's focus on climate change mistakenly neglects the fact that, "*The terrestrial climate is simulated through time as state magnitudes, not as anomalies*" *cf.* revised manuscript lines 704-705. This is a critically important point (Section 2.4.3), is physically correct, and is part of the "*shoddy philosophizing*" that the reviewer derided (*cf.* item 8.5).

Correction of the reviewer's mistake is shown here: in the reviewer's notation, the correct argument is not about C[T], but rather is about $(C \pm u)[T]$, i.e., C[T] <u>plus its calibration</u> <u>uncertainty</u>, where $\pm u$ is the simulation uncertainty in climate state cloud cover at a given [T].

This $\pm u$ arises from poorly constrained parameters and model theory error. Stemming from

the model itself, it is therefore imposed upon, and present within, every single step of a simulation.

Every simulated C[T] $\pm u$ thus has an associated C_h[T] and C_l[T], where C_h = C+u and C_l = C-u. The C+u and C-u represent the uncertainty range within which the physically correct magnitude of C likely occurs.

The $\pm u$ is uncertainty due to systematic error, and is determined through calibration experiments. The mean value of C within the $\pm u$ range is therefore not the most probable value. The physically correct value of C can be anywhere within $\pm u$ (or, less likely, outside of it).

Therefore, the simulated magnitude "C" cannot be assumed physically correct, and cannot be projected forward absent its uncertainty bounds.

Following the reviewer's example, the derivative is then,

$$d(C \pm u)[T]/dT = (d(C_h)[T]/dT, d(C_l)[T]/dT)$$
 R1

In R1, the subscripts "h" and "l" represent the high and low limits of the $\pm u$ uncertainty range. Thus, $d(C_h)[T]/dT$ is the high end and $d(C_l)[T]/dT$ the low end of the uncertainty range, respectively.

Every simulated dC[T]/dT is thus an associated range of values rather than a single value, representing the upper and lower uncertainty bounds of valid simulation magnitudes of C. These express the uncertainty limits of the simulated climate change.

In a given simulation step, '*i*,' $C_h[T]_i$ and $C_i[T]_i$, represent the range of physical uncertainty in the magnitude of $C[T]_i$, and are implicitly projected forward in each step of a simulation.

At simulation step (*i*+1), new $(C_h[T])_{(i+1)}$ and $(C_l[T])_{(i+1)}$ uncertainties then condition the new calculated $(C[T])_{(i+1)}$ expectation value, which initiated from $(C_h[T])_i$ and $(C_l[T])_i$.

When calculating the new change, $(dC[T]/dT)_{(i+1)}$, the new $(dC_h[T]/dT)_{(i+1)}$ and $(dC_l[T]/dT)_{(i+1)}$ uncertainties in simulated climate change necessarily impose themselves.

Following from 2.2.3, the new uncertainties in $(dC_h[T]/dT)_{(i+1)}$ and $(dC_l[T]/dT)_{(i+1)}$ are calculated as $\sqrt{\sigma_i^2 + u_{(i+1)}^2} = \pm \sigma_{C_i,C_{(i+1)}}$ and $\pm \sigma_{C_i,C_{(i+1)}} > \pm \sigma_i$.

The result is an expanding cone of simulation uncertainty across the projection.

It is thus immediately evident that the uncertainty in the base state propagates directly into simulated climate change.

The direct interpretation of this outcome is that for every dT, the magnitude of dC is uncertain.

Therefore the d(LWCF)/dC will include a $\pm d(LWCF)/dC$ calibration uncertainty because every dC is properly d(C $\pm u$).

The same propagation of calibration error and the same expanding cone of uncertainty follow. The reviewer cannot escape this result. Nor can the field.

2.2.6 Summarizing item 2.2, the reviewer has:

- Mistakenly assumed linear response theory
- Misconstrued uncertainty as physical error
- Does not know that calibration uncertainty propagates into simulations
- Does not understand uncertainty at all
- Has badly over-simplified the problem.

Such mistaken thinking pervades this reviewer's comments, and renders them void of critical content.

2.3 The reviewer's argument also implicitly assumes that the fractional distribution of cloud type remains constant as T varies.

However, the empirically derived LWCF uncertainty of $\pm 4 \text{ Wm}^{-2}$ almost certainly includes a contribution from incorrectly simulated cloud type, as well as incorrect spatial extents and incorrect distributions of cloud cover.

- 3. I believe that this assumption is wrong. It is neither obvious nor common sense that this is true and I think the published literature shows the assumption to be wrong.
 - 3.1 Items 2.2.1 and 2.2.3-2.2.5 show the author did not, "assume that the error in C[T] is equal to the error in dC[T]/dT."

The author cannot be wrong in an assumption he did not make.

The author showed that the reviewer has mistaken calibration uncertainty for physical error. The author has now also shown that the uncertainty in anomaly $(dC[T]/dT)_{(i,i+1)} = \pm \sigma_{(i+1)} > \pm \sigma_i$.

It is obvious that model calibration uncertainty propagates into a result and thence into an anomaly.

3.2 Further, in the round 1 response, the author repeatedly showed that the published literature ignores physical uncertainty in assessing climate simulations.

For example, the neglect of physical uncertainty was shown to be starkly evident in [*John and Soden*, 2007], the reviewer's own authority. Their Figure 2 includes no physical uncertainty bars.

This fact is evidently lost on the reviewer, despite that it was pointed out in author round 1 response item 1.1.

When uncertainty due to physical error is included, none of the points in [*John and Soden*, 2007] Figure 2 has any physical meaning at all. This, too, was established in the author's round 1 response. The lack of physical meaning deprives the reviewer's argument of any force.

- 3.3 Explicitly: it does not matter to physical uncertainty propagation if all models produce the same result, when that same result is physically meaningless.
- 4. As I mentioned in my original review, John and Soden (2007) showed that, despite large differences in the base state, all of the models predict the same water vapor feedback in other words, d(H2O[T])/dT is the same throughout the model ensemble despite significant differences in H2O[T].
 - 4.1 The reviewer has ignored the author's complete disproof of his contentions about the meaning of [*John and Soden*, 2007].
 - 4.2 To the contrary, the author showed in his round 1 response that [*John and Soden*, 2007] does not show all models predict the same water vapor feedback. Figure 2 merely shows that inter-model results are linearly correlated, not identical of magnitude.
 - 4.3 The author showed that none of the points in [*John and Soden*, 2007] Figure 2 is known to be physically correct.
 - 4.4 The author also showed that the GISS and CRU temperature measurements, used to judge Figure 2, are so uncertain that their error bars extend beyond the margins of the plots.

Thus, the physical accuracy of the simulation points in Figure 2 cannot even be evaluated. They provide no assurance at all of the physical accuracy of simulated climate change.

The reviewer has ignored these demonstrations and has addressed none of them. He has merely reiterated his claim right into the face of its analytical disproof.

4.5 The author demonstrated that [*John and Soden*, 2007] has nothing to say about the physical accuracy of a simulated change in climate.

This demonstration plus the reviewer's comment 4 response indicate that the reviewer clearly does not understand physical error analysis.

- 5.1 And Dessler (2013) showed that the models all produce similar cloud feedbacks, despite the large differences in the cloud fields documented in this paper.
 - 5.1 Once again, the reviewer has reiterated his claim in the face of the author's round 1 analytical disproof of it.

In round 1 response item 1.4, the author showed that [Dessler, 2013] used a modelderived reanalysis product that entrains a circular logic falsely validating model output. [Dessler, 2013] also:

- unjustifiably assumed the validity of linear response over extended times,
- ignored the impact on uncertainty of the incorrect partitioning of energetic flux revealed in the ERA-Interim and MERRA calculations,
- and applied a kernel calculation likely beyond its relevance.

The reviewer has ignored all this and has again uncritically cited [Dessler, 2013].

- 5.2 Both of these papers tell me that errors in the base state are not necessarily the same as errors in the derivative.
 - 5.2.1 The reviewer has ignored the author's demonstration that neither of those papers provide the evidence the reviewer has asserted.

The author regrets to observe, therefore, that the reviewer's critical perception is inadequate.

In round 1, the author demonstrated that neither paper says anything about physical accuracy or physical error. How was such an obvious point beyond the reviewer's grasp?

The reviewer has not shown that the author's analysis of these papers was incorrect.

5.2.2 In item 5.2, the reviewer is implying that, "*not necessarily the same errors*" = no residual error. The implied logic is incoherent.

For example, if the errors in the derivative are "*not necessarily the same*," why could they not be larger? The reviewer does not say.

The author finds this reviewer notably careless in considering model systematic error and its resultant calibration uncertainty.

- 5.2.3 In items 2.2.3 2.2.5 above, the author showed that the uncertainties in simulated magnitudes propagate into the derivatives.
- 6.1 The response from Frank to the John and Soden paper is extremely long and detailed, but seems to misunderstand my point. I am not arguing that the models are correct, only that the models show that the errors in the base state are not the same as the errors in the derivative.
 - 6.1.1 The author explicitly showed, *and stated*, that [*John and Soden*, 2007] did not assess error at all. It assessed only precision. The reviewer has evidently failed to grasp this absolutely fundamental point, rendering his 6.1 irrelevant.
 - 6.1.2 Contrary to the reviewer's 6.1, the author immediately addressed the reviewer's concerns. Thus, the reviewer began his round 1 review with the claim that [John and Soden, 2007] and [Dessler, 2013] showed, "that, despite large differences in the cloud fields, the change in clouds as the climate warms is basically the same. ... [and therefore that] taking an error in the base state and assuming that error translates into the error in the climate

response is unsupported by previously published analyses."

Round 1 response item 1.1 immediately and accurately began with, "[John and Soden, 2007] does not validate that the change in climate fields is accurately modeled." I.e., the author wrote directly to the reviewer's claims; first taking [John and Soden, 2007], and then [Dessler, 2013].

Thus, the author did not misunderstand the reviewer's point.

The author's analytical response went on to falsify the reviewer's contentions.

The author refuted the reviewer by showing that [John and Soden, 2007] is strictly concerned with inter-model comparisons, i.e., precision. [John and Soden, 2007] has nothing to say about accuracy or physical error.

Therefore [John and Soden, 2007] does nothing to demonstrate that the simulated change in climate is physically accurate.

It is clear, however, that this falsification was lost on the reviewer.

6.1.3 The reviewer is again implying that the statement, "*errors in the base state are not the same as the errors in the derivative*" is identical to 'no residual error.'

This same incoherent logic is exactly what the reviewer attempted to establish in round 1, when he wrote,

"John and Soden showed that, despite the biases in the water vapor fields, the **change** in water vapor in response to warming is nearly identical among the GCMs, meaning that the water vapor feedback is nearly identical. ... [Dessler, A. E. (2013)] means that, despite large differences in the cloud fields, the change in clouds as the climate warms is basically the same."

The reviewer thus implied both previously and again here that equivalence of simulated change among models is identical to no model error.

6.1.4 The reviewer has also again ignored the critical distinction between accuracy and precision. He has done so despite that the author extensively and repeatedly discussed and established this very distinction.

The conclusion of **no** physical error <u>does not follow</u> from a similarity of simulated change among climate models.

The reviewer's logic is incoherent, it flies into the face of scientific practice, and it is analytically false.

6.2 The Dessler paper doesn't explicitly make this point, but it does show that the feedbacks are generally similar between the models - there are some differences, of course, but these

differences are much smaller than the suggested error here.

- 6.2.1 The analytical and evidential weakness of Dessler, 2013 was thoroughly exposed in round 1 response item 1.4 and is again described in item 5.1 above.
- 6.2.2.1 Further, the reviewer is again confusing physical error and statistical uncertainty. The derived ±15 C centennial uncertainty in projected air temperature is not error. It is an uncertainty statistic.

The meaning of uncertainty was discussed exhaustively in SI Section 10, "*The Meaning of Uncertainty in Model Projections*." However the reviewer evidently ignored this discussion.

- 6.2.2.2 Additionally, the reviewer has again supposed that similarity among models is identical to simulation accuracy. It is not. Such similarity is mere precision. This point was repeatedly stressed and is central to the analysis. Nevertheless, the reviewer has never addressed it.
- 6.2.3 In any case, the author went ahead and tested the reviewer's suggestion that in Dessler, 2013, "[the long wave cloud] feedback differences are much smaller than the suggested error here."

Figure R1 below shows long wave cloud radiative feedback data digitized from Figure 6 (left) of Dessler, 2013. Note that the error is large with respect to observations.

As a first approach, the global observed and simulated LWC feedbacks were determined by integrating the areas under the curves of Figure R1. The globally integrated MERRA global longwave cloud radiative (LWR) feedback was 45.5 Wm⁻²K⁻¹, the simulated global cloud LWR feedback was 84.1 Wm⁻²K⁻¹, and the model error was 38.7 Wm⁻²K⁻¹. Thus, the magnitude of the model error is fully 85% as large as the MERRA global cloud LWR feedback itself. This is hardly the "*much smaller*" difference the reviewer has supposed.



Figure R1 Legend extracted from Dessler, 2013, Figure 6: "*The zonal average cloud feedbacks: (of) longwave ... radiation.*" MERRA observation is the solid red line, the model ensemble average is the solid black line, and the blue line is model error (ensemble minus MERRA). Negative numbers are South Latitude.

The global root-mean-square model ensemble error was then calculated across degrees of

latitude as
$$\pm \sigma_{\varepsilon} = \sqrt{\frac{\sum_{i=-90}^{90} (ensemble \ mean - MERRA)^2}{181}}$$
, where $\pm \sigma_{\varepsilon}$ is the uncertainty of the ensemble average with respect to observation in units of Wm⁻²K⁻¹ over each degree of latitude.

This RMS error of the ensemble mean is $\pm \sigma_{\epsilon} = \pm 1.5 \text{ Wm}^{-2}\text{K}^{-1}$ This value represents the global average simulation feedback uncertainty per degree latitude due to physical error; a measure of model accuracy (although this designation would be strictly true only if models were capable of unique solutions, which they are not).

However, a full accounting requires including the uncertainty due to the simulation spread of the individual models themselves, irrespective of the uncertainty due to physical error. This was estimated by taking pairs of points representing the upper and lower bounds of the one-standard deviation of the model spread in Dessler, 2013, Figure 6 (left). The global average of model spread was estimated to be ± 1.4 Wm⁻²K⁻¹. This second uncertainty represents model precision.

The total uncertainty in model long wave cloud feedback is the quadratic sum of the accuracy and the precision = $sqrt[(1.5)^2 + (1.4)^2] = \pm 2.0 \text{ Wm}^{-2}\text{K}^{-1}$.

Dessler, 2013 Table 1 reported average ± 0.36 Wm⁻²K⁻¹ as the simulation uncertainty standard

deviation for longwave cloud feedback. It is difficult to understand how this small value derived from the very wide standard deviation shading of Figure 6, left. The standard deviation of model spread is at least $\pm 1 \text{ Wm}^{-2}\text{K}^{-1}$ from 15° S to 45° N.

In any case, the combined long wave cloud feedback simulation uncertainty is some 4× larger than the empirical magnitude of the feedback itself.

Therefore, Dessler, 2013 does not support the reviewer's claim of a small uncertainty in projected air temperature.

- 7.1 Much of the author response is also based on the fact that the papers I cited are mainly model studies and don't explicitly validate the response in the models.
 - 7.1 The reviewer is correct, and has here and now falsified his own first round review, and on the same grounds as did the author: model inter-comparisons provide no indication of physical accuracy.

The model studies the reviewer originally offered as evidence of accuracy, in fact demonstrate only precision.

In comment 7.1, the reviewer has now agreed with that assessment, and contradicted his first round review. His initial case for rejection is now vacated.

- 7.2 If one is interested in specific validations of the model response, then there are many many papers that have estimated climate sensitivity, and they generally cover a range of 1.54.5[degree sign]C for doubled CO₂, in reasonably good agreement with the model ensemble.
 - 7.2.1 Any estimated climate sensitivity (ECS) to doubled CO_2 must include the same assumptions about the impact of CO_2 forcing as are present in climate models.

The author tested this idea by examining a number of papers providing estimates of ECS, i.e., [*Danabasoglu and Gent*, 2009; *Gregory et al.*, 2004; *Hansen et al.*, 1984; *Hegerl et al.*, 2006; *Xin-Yu et al.*, 2013]. Every single one of them required the use of climate models to obtain an estimate of climate sensitivity. This conclusion includes the empirical approaches of Hansen, et al., 1984.

7.2.2 The reviewer is invoking climate models to validate climate models. The reviewer's argument is logically and irremediably circular.

The [*Hegerl et al.*, 2006] paper, by the way, uses climate models plus paleo-temperature reconstructions to estimate ECS, assigning a ± 0.1 C accuracy to a physically meaningless "proxy" time series [*Frank*, 2015]. Such nonsense clearly exemplifies the critical failures plaguing the field.

7.2.3 An accurate estimate of climate sensitivity must derive from a relatively complete and accurate physical theory of climate. How, otherwise, is an accurate sensitivity to be calculated? The reviewer's argument rests upon the existence of such a physical theory. However, that theory does not exist.

- 7.3 It also means that the spread in the models is a reasonable estimate of the uncertainty in *future temperatures.*
 - 7.3 The reviewer first supposes an accurate estimate in the absence of an accurate physical theory.

The reviewer then supposes calculations based upon shared assumptions produce independent estimates. The reviewer finally supposes that comparability of the spread from these correlated estimates is a measure of accuracy.

That such reasoning can come from a scientist should be incredible, but apparently is not. That it should be accepted as a critical review should be equally incredible, but apparently is not.

7.4 This is yet another reason I suspect the results of this paper are entirely wrong.

7.4 Items 7.2.1 - 7.3 show the reviewer's reasoning to be incorrect and misguided throughout.

- 7.5 Let me be clear: there certainly ARE errors in the response derivatives in the models after all, equilibrium climate sensitivity does vary by a factor of about 2 in the models. But the uncertainty implied by this is much smaller than the uncertainty this author derives.
 - 7.5.1 The uncertainty the author derived is never taken into account in estimates of the physical uncertainty in ECS. The uncertainty from LWCF error alone, not to mention other large-magnitude uncertainties [*Rowlands et al.*, 2012; *Sanderson and Knutti*, 2012; *Soon et al.*, 2001; *Stephens et al.*, 2012], are never propagated into simulated air temperature.

The reviewer's range is merely the expectation value spread of various models that deploy a common theoretical structure and exhibit a common systematic error. These are simply measures of model precision. The reviewer himself admitted this in item 7.1.

7.5.2 The factor of two estimate does not reflect physical accuracy, and in no way provides a true indication of the uncertainty in the value of ECS. A comparison of the spread of ECS expectation values with the magnitude of a calibration uncertainty propagated through a projection, so as to dismiss the latter, is analytically incongruous. The reviewer's argument is entirely spurious.

The LWCF error alone makes climate model air temperature simulations physically meaningless. Thus the estimates of ECS are also meaningless. The reviewer's position is insupportable.

- 7.6.1 The author did address a few of my many other comments, but others were not adequately addressed.
 - 7.6.1 The above is starkly ironic. The author confronted and resolved every single point in the round 1 review, and in detail. The reviewer's comment 7.6.1 is doubly ironic because the

"*problem that remains*" (7.6.2, below), was explicitly resolved in the round 1 response, in items 4.3 and 4.4.

This reviewer utterly failed to confront the author's falsifications of his central claims regarding [*John and Soden*, 2007] and [*Dessler*, 2013] and went on to suggest it is the author who is remiss. He merely opined that he found the author's analysis unconvincing, which apparently suffices as a refutation.

- 7.6.2 One problem that remains is the estimate of CO2 fraction of the greenhouse effect. In my original review, I pointed out that the author assumes clouds have zero greenhouse effect (in his response, he protested that he didn't, but he should re-read lines 187+188 to see that he indeed does).
 - 7.6.2 The author did more than protest. He definitively showed the reviewer was factually incorrect, by quoting directly from the manuscript.

The reviewer here has only revealed a continuing inability to discern the obvious after two ostensibly critical appraisals. The quoted text below shows that cloud cover and its greenhouse effect is explicitly included in the manuscript analysis.

Manuscript lines 187-188, through 189, state, "The analysis begins in this section with an estimate of the fraction of the terrestrial greenhouse temperature produced by the wve forcing of CO₂, as relevant to general circulation climate models."

The paper of [*Manabe and Wetherald*, 1967] was used to make this estimate. [*Manabe and Wetherald*, 1967] specifically state that their green house calculations **include the greenhouse effects of cloud cover**.

Section 4c "Carbon dioxide," of [Manabe and Wetherald, 1967] begins with,

"As we mentioned in the introduction, Möller (1963) discussed the influence of the change in CO_2 content in the atmosphere with a given value of relative humidity on the temperature of earth's surface. ... In order to re-examine Möller's computation by use of the present scheme of computing radiative transfer, the net upward long-wave radiation into the atmosphere with given distributions of relative humidity was computed for various distributions of surface temperature, as shown in Fig. 9. The vertical distribution of cloudiness ... have already been specified in Section 2b. (underline added)"

[Manabe and Wetherald, 1967] Table 4 "Equilibrium temperature of the earth's surface (°K) and the CO_2 content of the atmosphere" likewise so indicates, with tabulated data for "Average cloudiness" and "Clear. The points in manuscript Figure 1, panel b, were taken from their Table 4.

Cloud cover and its greenhouse effect are clearly included in the calculations of [*Manabe and Wetherald*, 1967]. The results of those calculations compose the blue line in manuscript Figure 1, panel b. The legend reports the blue line conveys the air temperature under "*cloud cover*." The equation deriving the fractional impact of CO2 explicitly includes

the impact of cloud cover.

Despite all this -- the explicit inclusion of cloud cover in the [*Manabe and Wetherald*, 1967] study and the specific inclusion of the cloud cover effect in the manuscript -- the reviewer nevertheless claims, "*the author assumes clouds have zero greenhouse effect*." Even a casual examination of the manuscript disproves the reviewer's contention.

7.6.3 In so doing, the author arrives at CO2's fraction of the GHE is 42%.

This is larger than previously published estimates using the GISS climate model, which gets 20%. In his response, the author says that we shouldn't believe the 20% number because it comes from a single climate model, and other models might disagree.

7.6.3.1 The reviewer has misstated the both the author's response and the original assessment. The response pointed to the author's assessment in the manuscript, which the reviewer might have thought to consult. That assessment demonstrated the disparity of fractional CO_2 impact among climate models.

The author's response also explicitly pointed out the CO₂ sensitivity fraction among models ranged between 0.37 and 0.9. This is not that, "*other models might disagree*" as the reviewer has it (author bold). It is that they **do** disagree.

The manuscript demonstration (Figure 2 and text) showed the obvious disparity of the GISS model projection with the air temperature projections of all other climate models.

This disparity necessarily demonstrates that the fractional impact of CO_2 varies widely among models. The value derived from the GISS model has no special standing.

The same disagreement among the CMIP5 versions is evident in the RCP scenario projections (manuscript Figure 3 and Figure 4)

The reviewer's argument in 7.6.3 is thus both false and misapprehended.

- 7.6.3.2 The reviewer might have noted that the GISS models have an average $f_{CO_2} = 0.425 \pm 0.03$ (SI Tables S1-S4), i.e., the value the reviewer disputes.
- 7.6.4 So what does the author do? He bases his estimate on a 1967 paper by Manabe and Wetherald, a single 50-year old study that uses a single, radiative-convective model.

7.6.4.1 The author had searched the literature, and took the model-derived $[CO_2]_{atm}$ vs. temperature relationship from the only paper he could find (i.e., Manabe and Wetherald) that explicitly presented it.

The publication URL, available here, http://tinyurl.com/y73f9omv, shows that the five most recent citations to [*Manabe and Wetherald*, 1967] were 11 May 2017, 1 July 2017, 17 July 2017, 17 May 2017, and 10 May 2017, at this writing.

This is evidence that, contrary to the reviewer's derisive dismissal, this, "50-year old study

that uses a single, radiative-convective model" remains highly relevant right up through 2017.

Further, [Charnock and Shine, 1995] pointed out years ago that, "results from onedimensional (height) global mean radiative-convective models like ours are not inconsistent with the global means of results from three-dimensional (latitude, longitude, height) models."

Still further, manuscript Figure 2 and SI Figure S1-1 show beyond any rational dispute that $f_{CO2} = 0.42$ does an extremely good job emulating GCM air temperature projections. The reviewer seems impervious even to visually obvious demonstrations.

7.6.4.2 The reviewer's grammatical construction in 7.6.2-7.6.3 implies that the author consulted [*Manabe and Wetherald*, 1967] <u>in order</u> to refute the GISS analysis. It is difficult to reconstruct the reviewer's thinking here because the manuscript nowhere supports this supposition.

In fact, the author took data from [*Manabe and Wetherald*, 1967] Table 4 to enable the Figure 1 analysis, well before ever examining the GISS 20% in context.

- 7.6.5 I agree that one must look skeptically at a single GCM result, and an average of the ensemble of GCMs would clearly be preferable, but the Manabe and Wetherald paper should not be taken as more accurate than the far more modern GISS model analysis.
 - 7.6.5.1 The [*Manabe and Wetherald*, 1967] paper was not presented as more accurate. The reviewer here evidences a complete misunderstanding of the analysis around manuscript Figure 2.

That analysis merely showed that the disparate model projections indicated the nondefinitive nature of the GISS 20%. This is not hard to understand. Nevertheless, the reviewer has successfully failed to understand.

- 7.6.5.2 Supporting Information Tables S1 through S4 show that the GISS model uniformly indicates a fraction near 0.42, all the way through the CMIP5 GISS-e2-r-pl version.
- 7.6.6 That's particularly true given that an answer of zero is physically unreasonable: we KNOW that clouds cover a large fraction of the planet and they have a large impact on top-ofatmosphere long wave flux. There may be uncertainty in the value, but zero seems like a poor choice.

7.6.6 The reviewer's comment here is truly incredible.

Section 2.1.3 explicitly discusses cloud fraction. The discussion around Figure 1 explicitly discusses the use of cloud fraction in deriving the 0.42 value of f_{CO2} . The first term of the equation in line 304 explicitly includes the global 0.667 cloud fraction.

Figure 1, panel b includes the cloud fraction line from [*Manabe and Wetherald*, 1967]. The Legend points the reader to the cloud cover data.

Somehow the reviewer missed all of this while composing the review. A mere inspection of Figure 2 and text would have disabused the reviewer of his mistaken view.

There is **no** author assumption of zero cloud cover or zero cloud GH effect. Cloud fraction was included throughout.

If anything, comment 7.6.6 indicates the reviewer did not critically read the manuscript before providing this not-so-critical review.

- 7.6.7 In fact, there's a weird contradiction here: this entire paper is based on propagating the LW error of clouds. Yet the fact that the LW effect of clouds is important to the energy budget contradicts the assumption that the GHE of clouds is zero.
 - 7.6.7 Wrong again. As demonstrated in 7.6.1-7.6.6, there is no such assumption. It should be clear by now that this reviewer has been extraordinarily careless in review.
- 7.6.8 There remain other major problems in the paper, but documenting all of them would be a substantial investment in energy that I don't think the paper merits.
 - 7.6.8 To this point, not one of the reviewer's criticisms has survived critical examination. The review thus far lacks any analytical merit.
- 8. Other comments:
- 8.1 Comparing AMIP models to CMIP models is not a fair comparison (around line 506).
 - 8.1 There is no such comparison anywhere near line 506. The manuscript there introduces CMIP5 cloud error.
- 8.2 If this paper is published in anything close to its present form, which I don't recommend, I would insist it be substantially shortened. It reads like a Masters Thesis, which includes a lot of tutorial info, and I suspect it could be cut to about one-third of its present length by removing that which is well known in the field and other completely unnecessary text (e.g., lines 132-180).
 - 8.2 Lines 132-180 were added because previous reviewers complained that the Introduction was insufficiently explanatory. Much of the pedagogical discussion in the manuscript was added because reviewers, including the present reviewer, had clear problems understanding physical error analysis and uncertainty propagation.

The Supporting Information was meant to provide more information to readers about the meaning and impact of uncertainty. However, neither the present reviewer nor any other reviewer in the author's experience has given any evidence of having consulted it.

8.3 The paper focuses on LWCF error. However, the radiative error will tend to be offset by a compensating error in the SW. The author needs to discuss and estimate the magnitude of this compensation.

- 8.3.1 The author did this analysis in round 1, for reviewer #1. There, it was shown that SW makes little contribution to atmospheric thermal energy flux. There is no need to reproduce that analysis here. The present reviewer can request that analysis if so desired.
- 8.3.2.1 However, the reviewer should know that compensating errors do nothing to remove uncertainty.

The statistical reason is that sources of uncertainty combine as the root-sum-square (RSS) meaning that the sign distinctions of errors are lost. Inversely correlated (compensating) errors just means that the overall RSS uncertainty remains about constant.

The physical reason for this is that fortuitously cancelling errors provide no information concerning the underlying physical system. When the physics is wrong, no confidence can be had that an expectation value conveys any physical information about the state of the system, e.g., about its location in phase-space. Physical ignorance remains undiminished no matter that a "right answer" is obtained fortuitously.

In a climate projection, ignorance of phase-space error increases with each simulation step. This condition is communicated by propagating forward the physical uncertainty.

The reviewer's claim is that fortuitous error cancellation in a calibration hindcast proves an error-free projection.

Such a claim is groundless. It flies right into the face of scientific practice and meaning.

- 8.3.2.2 Additionally, the reviewer makes this claim even though he must know that error cancellation is typically obtained by a tendentious tuning of parameters [*Collins et al.*, 2011; *Kiehl*, 2007; *Rasch*, 2012].
- 8.3.3 The LWCF uncertainty provided in [*Lauer and Hamilton*, 2013] followed from averaging 20 years of hindcast from an ensemble of 27 CMIP5 models. Each model must have produced offsetting LW and SW errors, making LWCF error larger or smaller in inverse relation to SW error. The ensemble error average automatically included LW and SW error compensation, making the ±4 Wm⁻²/year statistical average LWCF error a true measure of the net uncertainty in tropospheric thermal energy flux, even in the presence of SW error.
- 8.3.4 Just to add, a "correct" magnitude produced by compensating error provides no physical understanding of mechanism (*cf.* 8.3.2.1, above). The uncertainty attending a fortuitous result conveys the actual persistence of ignorance.
- 8.4 Confusions abound in this paper. Sect. 2.4.3 is a hard-to-follow discussion about "differencing from a base state" and how that doesn't remove errors. On the other hand, the author put into his response to me (Fig. R2) a plot showing that the models have large differences in the absolute temperature of their simulated climate, but subtracting off the biases leads to quite good agreement in the temperature trend over the 20th century.

8.4.1 Agreement among models does not indicate removal of physical error. Such agreement

merely shows similarity of model structure.

The reviewer's confusion on this point is made repeatedly above in terms of the distinction between precision and accuracy, and was made throughout the manuscript.

Nevertheless, the reviewer has failed to grasp this critically important concept, though it is very basic to methodological practice in all the physical sciences. The confusion is therefore the reviewer's.

- 8.4.2 "*Subtracting off biases*," i.e., calculation of anomalies, leads to good agreement of the 20th century temperature trends because all the models have been tuned to reproduce the 20th century temperature trend [*Kiehl*, 2007]. This is obvious, published, well known, and tendentious.
- 8.5 This seems to contradict Sect. 2.4.3 and it points up another problem with the construction of the paper. There seems to be a lot of shoddy philosophizing in the paper (e.g., sect. 2.4.3) that contain no numbers or calculations and which are clearly contradicted by calculations made elsewhere.
 - 8.5.1 The "*shoddy philosophizing*" the reviewer derides includes that an 1850 base-state must have unknown errors, because the 1850 climate is observationally unknown.

It includes that models project climate as state magnitudes, rather than as anomalies.

It includes that theory-bias produces systematic errors, and that systematic errors do not average away.

It finally points to Sections of the Supporting Information that discuss all these points in detail. The discussion that includes all the analytical expressions for which the reviewer longs. It is apparent the reviewer never availed of this.

The above concepts are all very straightforward. That the reviewer finds them confusing merely indicates a lack of familiarity with methodological concepts common in the physical sciences and central to a scientifically valid evaluation of climate models.

The fact that such concepts have never been applied to climate models provides a complete explanation for the unjustifiable confidence infecting the field.

8.5.2 The fact that such concepts seem unknown to climate modelers provides a complete explanation for the lack of proper practice that has turned the field into a pseudo-science; a liberal art decorated with mathematics and imprisoned within its consensus narrative.

At this point it should be clear that Review #3 has no critical force.

References:

Charnock, H., and K. P. Shine (1995), How Cold Would We Get Under CO2-Less Skies?, Physics Today, 48(2), 78-80, doi: 10.1063/1.2807926.

Collins, M., B. B. Booth, B. Bhaskaran, G. Harris, J. Murphy, D. H. Sexton, and M. Webb (2011), Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles, Climate Dynamics, 36(9-10), 1737-1766, doi: 10.1007/s00382-010-0808-0.

Danabasoglu, G., and P. R. Gent (2009), Equilibrium Climate Sensitivity: Is It Accurate to Use a Slab Ocean Model?, Journal of Climate, 22(9), 2494-2499, doi: 10.1175/2008jcli2596.1.

Dessler, A. E. (2013), Observations of Climate Feedbacks over 2000–10 and Comparisons to Climate Models, Journal of Climate, 26(1), 333-342, doi: 10.1175/jcli-d-11-00640.1.

Frank, P. (2015), Negligence, Non-Science, and Consensus Climatology, Energy & Environment, 26(3), 391-416, doi: doi:10.1260/0958-305X.26.3.391.

Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31(3), n/a-n/a, doi: 10.1029/2003GL018747.

Hansen, J., A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lerner (1984), Climate Sensitivity: Analysis of Feedback Mechanisms, in Climate Processes and Climate Sensitivity, edited by J. E. Hansen and T. Takahashi, pp. 130-163, American Geophysical Union, Washington, D.C.

Hegerl, G. C., T. J. Crowley, W. T. Hyde, and D. J. Frame (2006), Climate sensitivity constrained by temperature reconstructions over the past seven centuries, Nature, 440(7087), 1029-1032, doi: 10.1038/nature04679.

Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations, J. Geophys. Res., 117(D14), D14105, doi: 10.1029/2011jd017237.

John, V. O., and B. J. Soden (2007), Temperature and humidity biases in global climate models and their impact on climate feedbacks, Geophys. Res. Lett., 34(18), L18704, doi: 10.1029/2007GL030429.

Kiehl, J. T. (2007), Twentieth century climate model response and climate sensitivity, Geophys. Res. Lett., 34(22), L22710, doi: 10.1029/2007gl031383.

Lauer, A., and K. Hamilton (2013), Simulating Clouds with Global Climate Models: A Comparison of CMIP5 Results with CMIP3 and Satellite Data, J. Climate, 26(11), 3823-3845, doi: 10.1175/jcli-d-12-00451.1.

Manabe, S., and R. T. Wetherald (1967), Thermal Equilibrium of the Atmosphere with a given Distribution of Relative Humidity, J. Atmos. Sci., 24(3), 241-259, doi: 10.1175/1520-0469(1967)024<0241:TEOTAW>2.0.CO;2.

Rasch, P. J. (Ed.) (2012), Climate Change Modeling Methodology: Selected Entries from the Encyclopedia of Sustainability Science and Technology, Springer Science+Business Media, New York.

Rowlands, D. J., et al. (2012), Broad range of 2050 warming from an observationally constrained large climate model ensemble, Nature Geosci, 5(4), 256-260, doi: 10.1038/ngeo1430.

Sanderson, B., and R. Knutti (2012), Climate Change Projections: Characterizing Uncertainty Using Climate Models, in Climate Change Modeling Methodology, edited by P. J. Rasch, pp. 235-259, Springer New York.

Soon, W., S. Baliunas, S. B. Idso, K. Y. Kondratyev, and E. S. Posmentier (2001), Modeling climatic effects of anthropogenic carbon dioxide emissions: unknowns and uncertainties, Climate Res., 18, 259-275.

Stephens, G. L., J. Li, M. Wild, C. A. Clayson, N. Loeb, S. Kato, T. L'Ecuyer, P. W. Stackhouse, M. Lebsock, and T. Andrews (2012), An update on Earth's energy balance in light of the latest global observations, Nature Geosci, 5(10), 691-696.

Xin-Yu, W., W. Shao-Wu, L. Yong, Z. Zong-Ci, and H. Jian-Bin (2013), Equilibrium Climate Sensitivity, Advances in Climate Change Research, 4(1), 69-72, doi: http://dx.doi.org/10.3724/SP.J.1248.2013.069.