Patrick Frank
Earth and Space Science Manuscript 2017EA000256
Response to Review #4

Summary response:

1. Physics is about causality, not plausibility; item 1.
2. "Plausible," unconditioned with a physically valid uncertainty envelope, has no physical meaning, item 1.
3. The systematic error propagation model follows widely published recommendations, items 2 and 3.1-3.4. Appropriate references have been added to the revised manuscript.
4. The reviewer misconstrued the ±4 Wm^{-2} LWCF uncertainty statistic as implying physical errors of sequentially opposite sign, item 3.5.
5. The centennial projection uncertainty is insensitive to the time over which calibration error is root-summed; item 4.
6. The reviewer has misconstrued the ±4 Wm^{-2} LWCF uncertainty statistic as a GHG forcing error; items 7 and 8.1.

The reviewer is quoted in italics below, followed by the indented author response. The reviewer's preamble is not addressed, except to observe that the response below belies the reviewer's conclusory recommendation.

1. Lines 60-16: The author claims that estimating the uncertainty in climate predictions by examining model variability relative to an ensemble mean is incorrect. However, in the absence of experimental observations (which occurs when making a forecast), then this approach is in fact one method for estimating the uncertainty due to the form of the models. This approach is known as "alternative plausible models" and is discussed in Morgan and Henrion (1990) and Cullen and Frey (1999).

> 1.1 Physics is about causality, not plausibility. Projections of future physical states are credible only when the physical models are known to produce unique and physically accurate solutions when tested against known physical observables. Climate models do not meet this standard.
>
> A reproduced recent air temperature trend, achieved by tuning a climate model, is neither a unique solution nor a demonstration of accuracy. "Plausibility" unconditioned by a physically valid uncertainty envelope is physically meaningless. This problem is illustrated below in Figure 1, taken from [*Lauer and Hamilton*, 2013; *Rowlands et al.*, 2012]. The discussion following the Figure applies equally to a hindcasted 20th century air temperature trend.

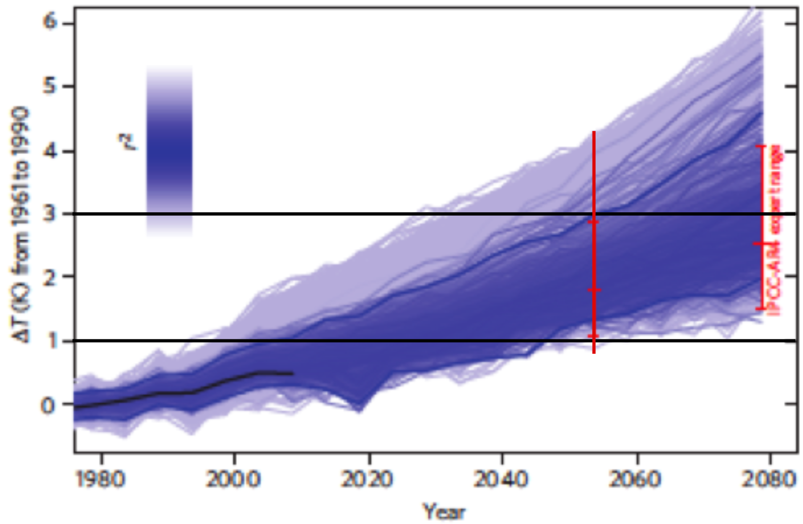


Figure 1. Original Legend: "*Evolution of uncertainties in reconstructed global-mean temperature projections under SRES A1B in the HadCM3L ensemble.*" The embedded black line in the original is the observed surface air temperature record. The horizontal black lines at 1 C and 3 C, and the vertical red line at year 2055, are author-added.

Figure 1 shows a set of perturbed physics projections wherein, "*a single model structure is used and perturbations are made to uncertain physical parameters within that structure...*" [*Collins et al.*, 2011]. That is, a perturbed physics experiment shows the variation in climate projections as model parameters are varied step-wise across their physical uncertainty width.

The horizontal black lines show the HADCM3L produces the same air temperature change for thousands of climate energy states.

For example, the upper black line shows that a constant 3 C increase in air temperature is projected for every single annual climate energy state between 2030-2080, depending on parameter set.

The identical logic applies to the vertical red line showing that the HADCM3L projects thousands of air temperatures for the single 2055 climate energy state. Every single annual climate energy state between 1976-2080 includes dozens of HADCM3L simulated air temperatures.

None of the different parameter sets producing these simulated temperatures is known to be any more physically correct than any other set. There is no way to decide which is physically correct among all the different choices of projected annual air temperature.

This set of examples shows that the HADCM3L cannot produce a unique solution to the problem of the climate energy state. Nor, by analogy, can any other advanced climate model.

No set of model parameters is known to be any more valid than any other set of model parameters. No projection is known to be physically correct, or any more physically correct than any other projection.

This means, for any given projection, the internal state of the model is not known to reveal anything about the underlying physical state of the true terrestrial climate. More simply, the model cannot tell us anything at all about the physically real climate, at the level of resolution of greenhouse gas forcing.

The same is necessarily true for any hindcasted climate energy state. Any hindcasted temperature trend is necessarily accompanied by very large uncertainties because the model is not known to correctly characterize the underlying physical state of the physically true climate.

Importantly, the identical criticism applies to the physical theory itself, as deployed within the models. It is neither known to be complete, nor entirely correct [*Su et al.*, 2013].

1.2 Regardless of any other considerations, evaluation of models by comparison of runs with an ensemble mean is a measure of precision only.

1.3 The accuracy of a model can only be determined by calibration against a known requisite physical observable or standard. Figure 1 above shows that no knowledge of model accuracy is presently available. Again, physics is about causality not plausibility. In this context, a judgment of "plausibility" then follows only upon a subjective acceptance of assumptions. In the absence of any possible strict empirical test, therefore, plausibility does not rise above the level of philosophy.

2. Line 66: The author claims that "Propagating physical errors through a model is standard in the physical sciences and yields a measure of predictive reliability." However, the only way you can propagate physical errors through a model is if they are random; if they are bias errors, then one should correct them and evaluate the model w/ the corrected inputs. If they are random, then you are really talking about uncertainty propagation (instead of physical error propagation).

2. The manuscript analysis follows directly from the prescription for propagation of error given in [*Bevington and Robinson*, 2003], which specifically includes empirically estimated standard deviations. The analysis also conforms with the recommendations of *The Evaluation of Measurement Data* provided by the International Bureau of Weights and Measures [*JCGM*, 100:2008].

The error propagation also follows [*Vasquez and Whiting*, 2006] who note that, "*When several sources of systematic errors are identified, [uncertainty] is suggested to be calculated as a mean of bias limits or additive correction factors.*" Their recommended equation (2) is the standard root-sum-square uncertainty as appears in the manuscript analysis.

The revised manuscript includes a new discussion of linear error propagation and provides supporting citations (lines 403-419, 599-601, 616-639 and 845-850).

The reported uncertainty statistic in longwave cloud forcing (LCF) from [*Lauer and Hamilton*, 2013] followed from a calibration experiment. Simulated cloud cover was evaluated against the observed cloud cover as standard. The errors from twenty-seven CMIP5 models were combined to yield the average 1σ LCF uncertainty, ±4 Wm^{-2}, representative of those models.

The manuscript further shows that LCF error is highly correlated among all tested climate models, and therefore arises from a systematic problem in the deployed theory. As do Bevington and Robinson, the [*JCGM*, 100:2008] recommends that the uncertainty from systematic error be treated using standard statistics, under "0.7 Recommendation INC-1" as follows:

"*3) The components in Category B should be characterized by quantities* u_j^2*, which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities* u_j^2 *may be treated like variances and the quantities* u_j *like standard deviations.*"

"4) The combined uncertainty should be characterized by the numerical value obtained by

applying the usual method for the combination of variances. The combined uncertainty and its components should be expressed in the form of "standard deviations"."

In 3) above, "Category B" components represent the uncertainty due to systematic error, *cf.* JCGM Section 3.3 and especially 3.3.3.

Section 4.3 in the JCGM specifies treatment of Category B uncertainties resulting from systematic error. Section 4.3.1, includes errors revealed by calibration experiments, i.e., the ±4 Wm^{-2} of [*Lauer and Hamilton*, 2013].

The combined standard uncertainty for both Type A and Type B uncertainty is specified in JCGM Section 5.1.2, equation 10 (author response eqn. 1):

"*The combined standard uncertainty $u_c(y)$ is the positive square root of the combined variance $u_c^2(y)$, which is given by*

$$"u_c^2(y) = \sum_{i=1}^{N} (\frac{df}{dx_i})^2 u^2(x_i) \qquad 1$$

"where f is the function [$Y = f(X_1, X_2, ..., X_N)$]. Each $u(x_i)$ is a standard uncertainty evaluated as described in 4.2 (Type A evaluation) or as in 4.3 (Type B evaluation). ... Equation (10) and its counterpart for correlated input quantities, Equation (13), both of which are based on a first-order Taylor series approximation of [$Y = f(X_1, X_2, ..., X_N)$], express what is termed in this Guide the law of propagation of uncertainty (see E.3.1 and E.3.2)."

JCGM equation 10 is recast to represent the combined uncertainty as JCGM 11a (author response eqn. 2):

$$u_c^2(y) = \sum_{i=1}^{N} u_i^2(y). \qquad 2$$

JCGM eqn. 11a is exactly the approach to propagation of uncertainty taken in the manuscript.

Further, [*Vasquez and Whiting*, 2006] discuss propagation of errors through nonlinear physical models. Their approach is entirely analogous to that from the JCGM, outlined above, and the manuscript analysis.

[*Vasquez and Whiting*, 2006] note that error propagation through nonlinear models is complicated and "*cannot be implemented successfully from an analytical standpoint.*"

The present analysis avoids this problem by propagating error through the linear model output, rather than through the model itself. This approach is validated by the demonstration that the nonlinear climate models invariably produce air temperatures that are linear with forcing.

The IPCC itself admits, in box 1.3 of [*Pyle et al.*, 2016], that climate model air temperature simulations are a linear function of forcing, i.e., $\Delta T_s = \lambda\Delta F$, where λ is climate sensitivity. This admission, in and of itself, validates a linear propagation of error.

3. *Line 87 (equation 1): The author gives the equation for uncertainty analysis used commonly in experimental measurements. However, there are three very important assumptions involved in the use of this equation. First, the equation should be linear over the range of uncertain inputs. Second, all of the uncertainties must be random in nature (i.e., they cannot be bias errors). Third, the uncertainty sources for the inputs must be uncorrelated. The author makes an argument for the first assumption (linearity),*

but the second and third assumptions are clearly violated. The error in the LWCF represents an unknown bias error. In addition, since the different uncertain inputs used in this manuscript come from a time sequence, they should clearly be correlated in time. That is, if the true error in LWCF in year 2 is +4 W/m2, it is unlikely that in year 3, the error in LWCF would be -4 W/m2.

3.1 Much of reviewer item 3 is already resolved in response item 2. First, linearity in model output is demonstrated.

3.2 As noted in item 2, systematic uncertainty is propagated in the manner of random uncertainty, noting the approximation inherent in the empirical u_j. This approach is also recommended for propagating non-random systematic error [*Garafolo and Daniels*, 2014; *Kacker et al.*, 2007; *Vasquez and Whiting*, 2006]. The approach is driven by the necessity of making a useful estimate of uncertainty in the real world of physical measurements, where statistical ideals are not often met. In this context, it is noted again that [*Lauer and Hamilton*, 2013] reported a calibration against real-world measurement data.

3.3 The ±4 Wm^{-2} is a root-mean-square uncertainty statistic derived from combining the calibration errors of twenty years of hindcast simulations from twenty-seven CMIP5 climate models, [*Lauer and Hamilton*, 2013]. As such, this uncertainty is representative of all CMIP5 models. As a multi-year averaged annual uncertainty, the LCF statistic is time-independent. That is, it is representative of the uncertainty in any simulation year of any current climate-futures projection. Time-series correlation has no meaning for a time-independent "±" rms uncertainty statistic.

Revised Section 2.3 now begins with new discussion of calibration error as applied to physical models. Revised Section 2.4 now begins with an extended discussion of the long wave cloud error statistic, its derivation, and its meaning.

3.4 It is not clear what the reviewer means by "unknown bias error." Model LWCF error arises from theory bias, i.e., the deployed physical theory is not correct. Incorrect physical theory produces an incorrect global cloud cover in a simulated climate [*Dolinar et al.*, 2015; *Jiang et al.*, 2012; *Pincus et al.*, 2008; *Su et al.*, 2013]. The source of this theory error within the model is unknown; were it known it could be corrected. However, the magnitude of this error is not unknown, but is known as made available by calibration against observables.

The simulated cloud cover includes regional positive and negative errors. It does not produce a global single-sign error offset. Regional errors means that air temperature (and thus convection) is incorrectly partitioned in a simulation. Further, the radiative effects of clear-sky and cloud-covered sky are incorrectly allocated across the global surface. The impact of this problem was discussed in manuscript lines 647-664 and 736-763.

3.5 The review's comment about time-wise correlation of error misconstrues the meaning of the ±4 Wm^{-2} uncertainty statistic. It does not imply errors of annually alternating signs, the reviewer's apparent perception. The statistic represents the annual average of uncertainty in the simulated tropospheric thermal flux intensity in each and every projection year. This point was discussed in manuscript section 2.4.2, line 554ff, again in lines 647-664, and in Section 3, line 729, and lines 736-763.

Sections 2.3 and 2.4 of the revised manuscript now provide a discussion of the meaning of this statistic. The derivational logic is now presented in new SI Section 6.2, along with the dimensional analysis. It is hoped these resolve the reviewer's concerns.

4. Pages 5-6: The use of the LWCF in each year as an independent, random, uncertain input appears to be *flawed. For example, what happens if the time sequence is broken down into smaller increments, say months instead of years? Does the uncertainty get much larger? What if it is in increments of 10 years instead of 1 year. Does the uncertainty in the final prediction get drastically smaller? If the answer to either of these questions is yes (which I believe it would be), then this approach is clearly flawed.*

4. If the time sequence is calculated over different time averages, then the average change in cloud cover scales. This is shown below for 20-year or monthly averages. The uncertainty in air temperature remains comparable throughout.

   [*Lauer and Hamilton*, 2013] describe their calculation of the LCF error statistic as, "*A measure of the performance of the CMIP model ensemble in reproducing observed mean cloud properties is obtained by calculating the differences in modeled (x^{mod}) and observed (x^{obs}) 20-yr means. These differences are then averaged over all N models in the CMIP3 or CMIP5 ensemble to calculate the multimodel ensemble mean bias Δ^{mm} which is defined at each grid point as*

$$\Delta^{mm} = \frac{1}{N}\sum_{i=1}^{N}(x_i^{\mathrm{mod}} - x^{obs}) \qquad (1)"$$

   Response Figure 2 below, taken from Figure 2 of Lauer and Hamilton, 2013, shows the CMIP3 and CMIP5 LCF error with the intensity bar.

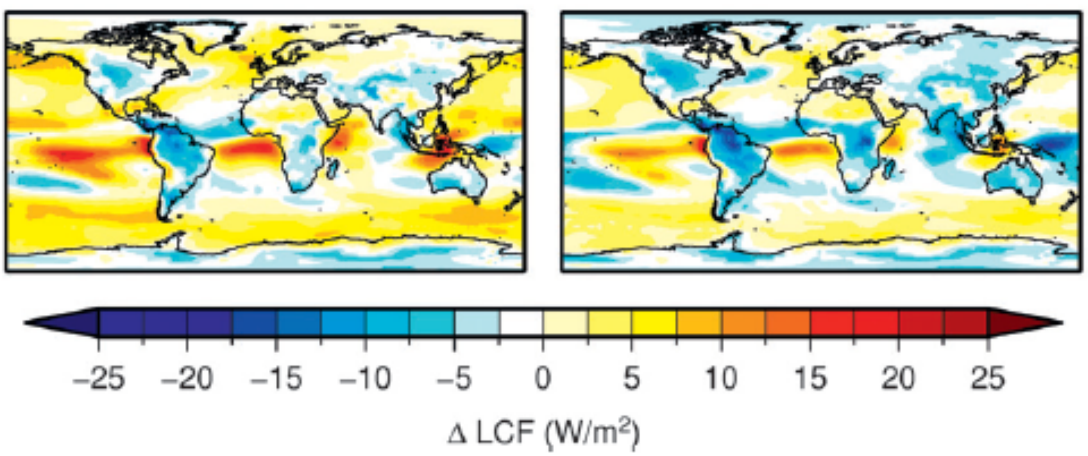


   Figure 2, original Legend: "*Differences in 20-yr annual averages of ... LCF from the (left) CMIP3 and [(right)] CMIP5 multimodel means compared with ... satellite observations.*"

   The LCF error means show regional positive and negative excursions and thus is not a single-sign offset or bias error. Cancellation of positive and negative sign error by addition misrepresents the accuracy of the calculation and produces a false precision.

   The multi-model annual rms error is calculated as $\pm\sigma_{LCF} = \sqrt{\frac{\sum_{i=1}^{26}(\overline{x}_i^{\mathrm{mod}} - \overline{x}^{obs})^2}{27}}$, where $\overline{x}_i^{\mathrm{mod}}, \overline{x}^{obs}$ are the 20-year mean modeled and mean observed cloud forcing, respectively.

   The annual average LCF error, ±4 Wm^{-2} reveals a 16 Wm^{-2} model annual average error variance. If recalculated as a 10-year block-average, rather than an annual average then, in equation (1) above, N = 2 rather than 20. In this case, the rms error requires a per-model mean error from a 10-times larger root. That is, it requires using the mean model error in Wm^{-2} summed across 10-year units.

The 10-year average variance is then 160 Wm^{-2}, yielding a 10-year LCF uncertainty σ =±12.6 $Wm^{-2}(10y)^{-1}$.

Manuscript uncertainty equations 7,8 yield the uncertainty across a 10-year step as (33K × 0.42 × 12.6 Wm^{-2})/33.321 Wm^{-2} = ±5.24 K, where 33.321 Wm^{-2} is the GHG forcing for the year 2000. After 10 such steps, the uncertainty in a centennial projection year 2100 is $\sqrt{10\times(5.24)^2}$ = ±16.6 K, i.e., an uncertainty entirely comparable to the manuscript value calculated in annual time-steps.

Likewise, for a monthly average LCF uncertainty N = 240 rather than 20, and yielding monthly LCF rmse = ±1.2 Wm^{-2}. The centennial uncertainty in air temperature across 1200 monthly time steps is ±17.3 K, again entirely comparable.

5. Regarding Figure 5 and the corresponding discussion around lines 432-435, the author makes the argument that since the TCF versus latitude curves all have similar shapes, that the different models "do not display random-like dispersions around the zero-error line." However, while there may be systematic error in TCF versus latitude, but the actual mean over the earth's area (which I believe would be much more relevant here) would be much less and may display a more random behavior.

5. Response Figure 2 above shows the CMIP5 20-year, 27 model, multimodel mean error in LCF, representing an average across a total of 540 simulation years and a 23× reduction in random error. The remaining errors must be systematic.

   As noted in item 4 above, cancellation by addition of regional-scale positive and negative errors hides the uncertainty and produces a false precision.

6. Lines 511-512: The author states "CMIP5 models were found to produce an annual average LWCF root-mean-squared error (rmse) = ±4 Wm^{-2}." The main point of this manuscript seems to revolve around this uncertainty magnitude. However, I suspect that it is possible that this large uncertainty arises from the fact that each one of the CMIP5 models is highly calibrated based on historical data. In some cases, the LWCF values may have been calibrated to better match the data. However, since it is accepted that the climate models are not perfect (i.e., that they contain significant model form uncertainty), then it is possible that the LWCF values, when examined across a suite of different models, may exhibit variations that are not consistent with the actual uncertainty in LWCF values. Stated differently, the uncertainties in LWCF are likely over-estimated due to the calibration of models with acknowledged model form errors.

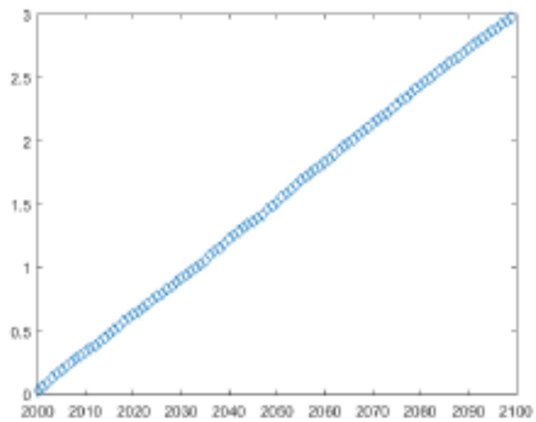
6. The reviewer's reasoning is difficult to understand. If models are tuned to match known LWCF, this greater matching reduces LWCF calibration errors. Then how can it be that LWCF errors are over-estimated?

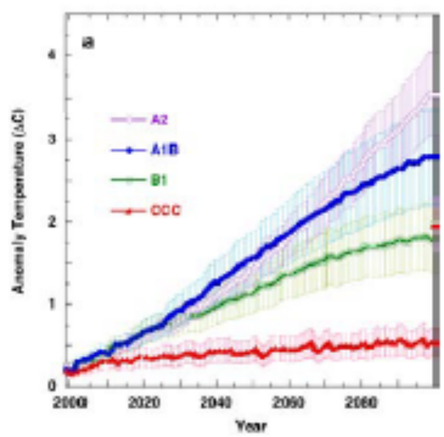
   That is, tuning to reproduce known LWCF should diminish the differences with observed LCWF. Model LWCF errors should then be under-estimated in a calibration experiment, not over-estimated.

   [*IPCC*, 2013] describes model tuning in box 9.1, Chapter 9. Typically, selected suites of parameters are adjusted so as to reproduce the radiative balance at the top of the atmosphere. Evidently, different groups tune different combinations of parameters to attain this balance, but do not fully report their choices. The CMIP5 experimental design can be consulted in [*Taylor et al.*, 2012].

   The LWCF uncertainty derived in [*Lauer and Hamilton*, 2013] represents 27 CMIP5 models. These models were tuned to climate observables and are representative of the models that are used to project climate. Their LWCF error is thus appropriate to a representative propagation of uncertainty.

7. In order to understand the author's proposed emulation model (equation 6, the Passive Warming Model, or PWM) better, I programmed up the model and applied it to predictions of year 2000 through 2100. Without knowing how the incremental change in greenhouse gas forcing (ΔF) values varied, I simply assumed a constant value. This value was chosen in order to match reasonably well to Figure 7 from the manuscript. The results for mean temperature anomaly are shown below (left) along with Figure 7 reproduced from the manuscript (right). In order to match the final temperature anomaly reasonably well, I had to assume an incremental change in greenhouse gas forcing ΔF of 0.07 W/m^2. However, the quoted uncertainty in this value is 4 W/m^2, which represents one standard deviation. It seems quite implausible to me that the average change in greenhouse gas forcing from year to year is 0.07 W/m^2, while the uncertainty in this number is 4 W/m^2. Something is not right here. Again, I suspect this uncertainty is much too large, as mentioned in item #7 above.





Reviewer's Simulation for Mean Global Temperature Anomaly Fig. 7 (from manuscript)

7. The reviewer apparently chose to model the SRES A2 scenario. For SRES A2, the 2000-2099 average annual change in forcing is 0.067 Wm^{-2}, so the reviewer's estimate is good.

However, the uncertainty in this value is not the reviewer's 4 Wm^{-2}, or rather ±4 Wm^{-2}. The ±4 Wm^{-2} is an uncertainty in the tropospheric thermal energy flux, consequent to model error. It is not an error in GHG forcing or in SRES scenario forcing. The A2 forcing uncertainty is zero, because the SRES forcings are assigned.

Rather, the ±4 Wm^{-2} is the uncertainty in the simulated tropospheric thermal energy flux arising from model LWCF error. GHG forcing becomes a part of the tropospheric thermal energy flux.

A large uncertainty in the simulated flux itself strongly conditions an ability to resolve the impact of the small perturbation to simulated tropospheric flux represented by GHG forcing.

This point is made clear throughout the manuscript. The concept of LWCF error as a tropospheric thermal energy flux error was introduced in manuscript Section 2.4.1 "*The magnitude of CMIP5 TCF global average atmospheric thermal energy flux error*."

Line 501 immediately under the heading states, "*CMIP5 TCF error entrains an error in simulations of tropospheric thermal energy flux.*" Line 510 points out that, "*LWCF represents the contribution made by clouds to the thermal radiation bath of the atmosphere.*"

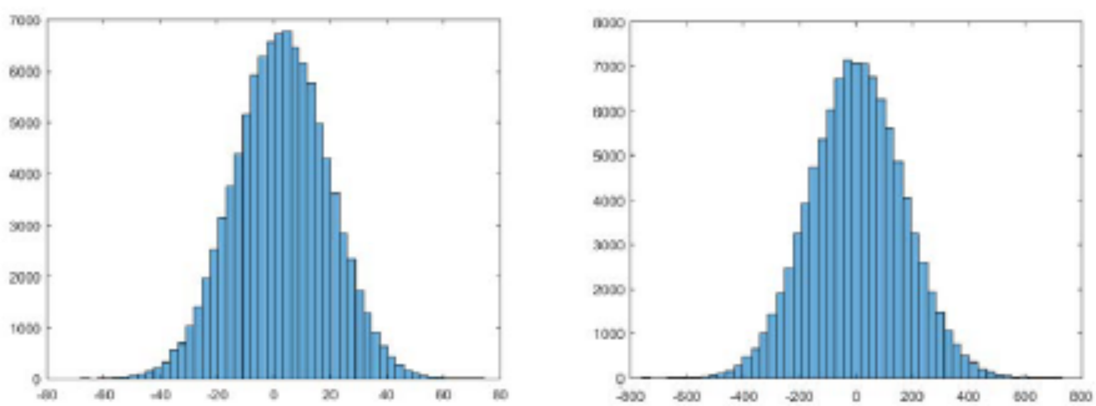
Line 518-523 again made this point, and concluded, "*simulations of the climatic response to changes in GHG atmospheric forcing are limited by ±4 Wm^{-2} of uncertainty in the magnitude of* *thermal energy flux within the troposphere.*"

That is, the impact of the SRES A2 0.07 Wm^{-2} annual increase in GHG forcing must be resolved against an annual uncertainty of magnitude ±4 Wm^{-2} in the simulated thermal energy flux of which GHG forcing becomes a part.

This is the meaning of manuscript line 533, "*In eqn. 7, $F_0+\Delta F_i$ represents the tropospheric GHG forcing, which is now conditioned by the uncertainty in simulated tropospheric thermal energy flux due to LWCF error*."

Further explanations of the significance of LCF error with respect to projection uncertainty are found in lines 542-552, 557-562, and 792-809.

8. As mentioned before, the author assumed that uncertainties in thermal flux (due to LWCF error) from year to year are independent, random variables. In order to test the effects of this assumption, I ran Monte-Carlo simulations (using 100,000 samples) using my rough approximation of the author's PWM. When assuming that the input uncertainties were independent, I found that the final temperature anomaly at year 2100 had a mean of 2.76 K and a standard deviation of 16.7 K, which is consistent with the result given in the manuscript. I then re-ran the simulations assuming that the uncertainties were perfectly correlated from year to year (i.e., using the same uncertainty sample each year). The result was that the year 2100 temperature anomaly had a mean of 2.78 K and a much larger standard deviation of 166 K. Histograms of these two simulations are shown below: uncorrelated (left) and perfectly correlated (right). While the true result is likely somewhere in between, it strains credibility to believe that the actual one standard deviation uncertainty in global temperature anomaly could be on the order of ±100 K and that the two standard deviation uncertainty (~95% probability) could be in the range ±200 K.



Independent, Random Uncertainties *Perfectly Correlated Random Uncertainties*

Year 2100 Temperature Anomaly Histograms

8.1 The reviewer is mistaken in the first sentence of item 8. Referencing the statistic as a "variable," as the reviewer does, implies the statistic is a physical quantity, which it is not. That is, representing an error statistic to be a variable entrains a very basic category mistake. It supposes a statistic is a physical magnitude.

The LWCF error statistic was not assumed to be an independent random variable. Rather, it is accepted as [*Lauer and Hamilton*, 2013] provided it, namely as a model calibration uncertainty.

The LWCF calibration error is an annualized average uncertainty statistic representative of CMIP5 climate models. It therefore and necessarily conditions each and every simulation year of a climate projection made using such models.

8.2 If a model LWCF error is time-wise autocorrelated, then a corrected 20-year mean error might be calculated using an effective time $(N_Y)' = 20\times[(1-r)/(1+r)]$, where "r" is the autocorrelation coefficient.

[*Lauer and Hamilton*, 2013] do not mention whether the tested model LWCF errors were time-wise autocorrelated. They only mention the correlation of simulated and observed LWCF, ranging from 0.70-0.92 for the tested CMIP5 models.

8.3 The reviewer's comment that "it strains credibility" amounts to an argument from personal incredulity, which has no strength in a scientific context. The meanings of 2σ uncertainties of ±33.4 K or ±200 K are identical, namely that the projection transmits no physical information about the future climate state.

According to [*Stephens et al.*, 2012], the measurement uncertainty in the terrestrial surface energy budget is ±17 Wm^{-2}. Thermal radiation from the warm surface is by far the greatest contributor to the tropospheric thermal energy flux [*Costa and Shine*, 2012; *Stephens et al.*, 2012; *Whitlock et al.*, 1995]. Thus, a fuller account of the tropospheric thermal flux uncertainty would include the contribution the uncertainty in the surface flux budget makes to the total uncertainty in the tropospheric energy flux budget.

[*Costa and Shine*, 2012] point out that only a tenth of the surface flux intensity emerges as outgoing LW radiation at the TOA. The rest of it (0.9) enters the troposphere before finally emerging at the TOA as part of the outgoing LW emission.

Known tropospheric thermal energy flux is a target of climate model simulations. Uncertainty in the target value imposes its uncertainty onto a simulated value. In that case, should not something like $0.9\times\pm17\ Wm^{-2} = \pm15\ Wm^{-2}$ be included in a fuller account of the uncertainty attached to simulated tropospheric thermal energy flux?

A tropospheric flux uncertainty of ±4Wm^{-2} and ±15 Wm^{-2} combined in quadrature is ±15.5 Wm^{-2}. Inclusion of this one additional source of uncertainty in simulated tropospheric thermal flux, alone, propagates across 100 years to a centennial 2σ uncertainty in projected air temperature of ±129 K; a value comparable to that which strained the reviewer's credibility. However, this addition does not tell us anything more than did the propagation of ±4 Wm^{-2}. Each of them says the same thing, which is that no physical information about the future climate state is available in the model projection.

A relatively comprehensive evaluation of climate model error was made some time ago, indicating combined errors of order 100 Wm^{-2} [*Soon et al.*, 2001]. Although a similar study has not been carried out on CMIP3 or CMIP5 models, much of the accounting remains relevant. That level of error indicates nearly a third of incoming solar energy is incorrectly partitioned among the climate sub-states. One can only imagine the magnitude of uncertainty in a projected climate, after 100 years of 100 Wm^{-2} uncertainty is propagated through a climate model simulation.

References:

Bevington, P. R., and D. K. Robinson (2003), Data Reduction and Error Analysis for the Physical Sciences, 3rd ed., McGraw-Hill, Boston.

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.

Costa, S. M. S., and K. P. Shine (2012), Outgoing Longwave Radiation due to Directly Transmitted Surface Emission, Journal of the Atmospheric Sciences, 69(6), 1865-1870, doi: 10.1175/jas-d-11-0248.1.

Dolinar, E. K., X. Dong, B. Xi, J. H. Jiang, and H. Su (2015), Evaluation of CMIP5 simulated clouds and TOA radiation budgets using NASA satellite observations, Climate Dynamics, 44(7-8), 2229-2247, doi: 10.1007/s00382-014-2158-9.

Garafolo, N. G., and C. C. Daniels (2014), Mass Point Leak Rate Technique with Uncertainty Analysis, Res. Nondestr. Eval., 25(2), 125-149, doi: 10.1080/09349847.2013.861953.

IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group 1 to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*Rep. 5*, 1535 pp, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

JCGM (100:2008), Evaluation of measurement data — Guide to the expression of uncertainty in measurement *Rep. Document produced by Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1).* Bureau International des Poids et Mesures, Sevres, France.

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.

Kacker, R., K.-D. Sommer, and R. Kessel (2007), Evolution of modern approaches to express uncertainty in measurement, Metrologia, 44(6), 513.

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.

Pincus, R., C. P. Batstone, R. J. P. Hofmann, K. E. Taylor, and P. J. Glecker (2008), Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models, J. Geophys. Res.: Atmospheres, 113(D14), D14209, doi: 10.1029/2007jd009334.

Pyle, J., et al. (2016), Chapter 1. Ozone and Climate: A Review of Interconnections, in Safeguarding the Ozone Layer and the Global Climate System: Issues Related to Hydrofluorocarbons and Perfluorocarbons, edited by T. G. Shepherd, S. Sicars, S. Solomon, G. J. M. Velders, D. P. Verdonik, R. T. Wickham, A. Woodcock, P. Wright and M. Yamabe, IPCC/TEAP Geneva.

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.

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.

Su, H., et al. (2013), Diagnosis of regime-dependent cloud simulation errors in CMIP5 models using "A-Train" satellite observations and reanalysis data, J. Geophys. Res.: Atmos., 118(7), 2762-2780, doi: 10.1029/2012jd018575.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An Overview of CMIP5 and the Experiment Design, BAMS, 93(4), 485-498, doi: 10.1175/bams-d-11-00094.1.

Vasquez, V. R., and W. B. Whiting (2006), Accounting for Both Random Errors and Systematic Errors in Uncertainty Propagation Analysis of Computer Models Involving Experimental Measurements with Monte Carlo Methods, Risk Analysis, 25(6), 1669-1681, doi: 10.1111/j.1539-6924.2005.00704.x.

Whitlock, C. H., et al. (1995), First Global WCRP Shortwave Surface Radiation Budget Dataset, Bulletin of the American Meteorological Society, 76(6), 905-922, doi: 10.1175/1520-0477(1995)076<0905:fgwssr>2.0.co;2.