Patrick Frank
Earth and Space Science Manuscript 2017EA000256
Response to Review #3

Summary response:

1. The reviewer misconstrued eqn. 6 to be about climate physics, rather than about the behavior of climate models; item 1.
2. The reviewer mistakenly assigned the ±4 Wm^{-2} of uncertainty in tropospheric thermal energy flux to GHG forcing; items 2.2 and 3.
3. The reviewer misconstrued statistical uncertainty as physical error; items 2.2, 2.3, 2.4, 3, and 4.
4. The reviewer has confused a mean of error with a mean of uncertainty; item 2.3.
5. The reviewer apparently does not understand propagation of error, item 4.
6. Throughout, the reviewer neglected the "±" sign when referencing the ±4 Wm^{-2}, i.e., it is always written as a positive sign 4 Wm^{-2}. This illustrates the reviewer's invariable misconstrual of statistical uncertainty as a physical error.

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

The review preamble is not addressed except to note that the reviewer is correct in observing that the manuscript analysis calls into question historical climate projections.

1. The theory is based on linear statistics; however, temperature power 4 is calculated for radiation. This cannot be linearized for 33 K temperature difference. This misunderstanding together with Eq. 6 may happen to represent its effect on radiation, but the theoretical basis is not solid.

> 1. The reviewer refers to the Stefan-Boltzmann equation. Nothing in the manuscript contradicts the S-B relationship. Nothing in the manuscript depends upon the S-B equation.
>
> Manuscript eqn.6 does not concern climate physics, or the Stefan-Boltzmann 4th power relationship. Equation 6 concerns the behavior climate models and the linear relation the models exhibit between simulated air temperature and GHG forcing.
>
> Manuscript Figures 2,3 4, 8 and 9 demonstrate this linear relationship, as do Supporting Information Figure S1, and Figure S3 through Figure S8.
>
> Further in box 1.3 of [*Pyle et al.*, 2016] the IPCC itself admits the linear relationship between radiative forcing and surface air temperature as $\Delta T_s = \lambda \Delta F$, where λ is a climate sensitivity that varies from model to model.
>
> The linear statistics of error propagation follow from the demonstration that models project air temperature linearly with forcing. The basis for linear statistics is obvious. Linear statistics is an empirical consequence, not an assumption.
>
> The reviewer has misunderstood the most basic point, namely that the analysis concerns

the behavior of climate models. It does not concern climate physics.

2. The "forcing" term used here is confusing. Sometimes it is for climatology, and sometimes for change.

2.1 it is difficult to see the reviewer's problem with the use of "forcing." Manuscript line 144 defined water-vapor-enhanced CO_2 forcing. This meaning is used throughout. The term has this standard usage everywhere in the manuscript.

2.2 The major problem here is the 4 m w-2 error is for total F0 but not dFi.

2.2 F_0 is the assigned greenhouse forcing of a simulation base-state, while dF_i is the annual change in GHG forcing designated within a chosen projection scenario.

In stark contrast, the ±4 Wm^{-2} derives from cloud feedback error; from a simulation error. It is not a GHG forcing error; it is no part of F_0, nor part of dF_i.

The ±4 Wm^{-2} is the mean annual uncertainty in simulated tropospheric thermal energy flux, which [*Lauer and Hamilton*, 2013] derived as $\pm\sigma_{LWCF} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sigma_{i_{LWCF}}^2}$, where subscript "*LWCF*" is longwave cloud forcing, *N* is the number of CMIP5 climate models in their study, and $\pm\sigma_{i_{LWCF}}$ is the mean annual uncertainty in longwave forcing for an individual CMIP5 climate model. The empirical $\pm\sigma_{i_{LWCF}}$ is obtained from calibration against 20 years of simulation hindcast.

That is, ±4 Wm^{-2} is an uncertainty in simulated tropospheric thermal flux that conditions the accuracy of the simulated effect of GHG forcing. Its status as a simulation conditional is why it must be entered into manuscript eqn. 6.

In other words, GHG forcing becomes part of the tropospheric thermal flux. Uncertainty in the simulated tropospheric thermal flux (±4 Wm^{-2}) impacts resolution of the thermal effect of GHG forcing. There is no escaping the problem of model resolution.

Once again, the ±4 Wm^{-2} is neither part of F_0, nor part of dF_i.

2.3 The annual error is annual mean of the error, but not error generated every year.

2.3 The ±4 Wm^{-2} is the annual mean of uncertainty, not the mean of error. A mean of error is $\mu_\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i$, where ε_i is some simulation error. The mean of uncertainty is $\pm\mu_u = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2}$. In a mean of error, errors of opposite sign cancel. Signs to not cancel in a mean of uncertainty.

[*Lauer and Hamilton*, 2013] reported a model calibration experiment. The ±4 Wm^{-2} is a 20-year mean uncertainty statistic derived from 27 CMIP5 climate models (540 model simulation years). This ±4 Wm^{-2} is representative of the annual uncertainty in simulated tropospheric thermal flux for any CMIP5 projection.

Understanding the difference between error and uncertainty is critical. Error is a physical quantity, e.g., simulated minus observed. Uncertainty is a statistic concerning the state of knowledge (certainty of result). It is propagated into a result when an error calculation is unavailable; that is, when the physically correct magnitude is not independently known.

Typically, uncertainty is derived from a calibration experiment, in which the accuracy of a model (or an experimental method) is determined by its ability to reproduce a known magnitude. The accuracy statistic is then imported into a model projection as an indicator of reliability of result, e.g., for a future climate where the physically true value is unknown.

When a sequential series of calculations is performed, each one of which entrains a known uncertainty, then the final uncertainty is the calibration uncertainty for each step propagated through the calculational sequence.[*Vasquez and Whiting*, 1998; 2006]

In a climate projection, error is produced in every simulation step. However, neither the sign nor the magnitude of error is known. The only known is the uncertainty.

Uncertainty necessarily increases with every simulation step, because the state of knowledge concerning the relative phase-space positions of the simulated climate and the physically true climate are unknown. The relative phase-space positions evolve in an unknown manner in simulation time.

When only uncertainty is available, the magnitude of error is unknown. The true but unknown physical error may be small even though the uncertainty is large (e.g., see [*JCGM*, 100:2008] "Note" page 6).

All of this was extensively discussed under manuscript Section 2.4.3, under Section 3, lines 725ff, as well as in Supporting Information Section 7 and most especially SI Section 10.2 "The meaning of predictive uncertainty".

A new paragraph has been added to Section 2.4.1 discussing the meaning of a calibration experiment.

2.4 The cumulative effect is fake. This is the reason for ~100 times overestimation of the error.

2.4 Propagation of uncertainty through a calculation is a standard of error analysis in the physical sciences [*Bevington and Robinson*, 2003; *JCGM*, 100:2008; *Taylor and Kuyatt.*, 1994]. It is properly applied to the expectation values of complex non-linear models [*Helton et al.*, 2010; *Roy and Oberkampf*, 2011; *Vasquez and Whiting*, 2006].

While the growth of uncertainty is often unfortunate to a conclusion, it is not "fake."

Response sections 2.3 and 2.4 follow from the reviewer mistake noted in response item 2.2, namely that the ±4 Wm^{-2} is an uncertainty in feedback. It is not part of F_0, nor part of dF_i

*3. The 4 w m-2 error itself is for F0, which means the error for GHG global warming forcing should be dF*12.1% but not F0*12.1%.*

3. The mistake represented by review item 3 was resolved in response item 2.2 above. However, the reviewer should have noted that the ±12.1% error in cloud cover conditioned forcing nowhere in the manuscript analysis; neither F_0 nor dF_i.

   Again, ±4 Wm^{-2} is the annual mean uncertainty in CMIP5 longwave cloud forcing (LWCF). It is not part of F_0, which is the base-year greenhouse gas forcing, nor part of dF_i, which is the annual change in greenhouse gas forcing.

   It is not clear at all how the reviewer has made the association of F_0 or dF_i with the rms LWCF error because neither association appears anywhere in the manuscript.

   As noted in response item 2.2, ±4 Wm^{-2} enters eqn. 6 because it represents the uncertainty in the simulated tropospheric energy thermal flux, of which GHG forcing becomes a part.

   The ±4 Wm^{-2} annual uncertainty width is then the background against which the average annual 0.035 Wm^{-2} GHG forcing increase must be resolved.

   As the background annual simulation uncertainty is about ±114 times larger than the annual GHG perturbation, it is quite clear that resolution of any GHG effect is impossible.

4. The conclusion is obviously wrong, since the annual error should have existed even without global warming. The climate models would go everywhere if this amount of error exists.

4. The reviewer is correct that the annual uncertainty, reflective of known model error, should exist apart from global warming.

   Climate models are known to make large energetic errors [*Collins et al.*, 2011; *Soon et al.*, 2001]. However, models are tuned to known observables, such as the TOA radiation balance. It is therefore not surprising that their outputs are both reasonable with respect to past observables and comparable among models.

   [*Kiehl*, 2007] has noted that tuning produces parameters with offsetting errors, which is why models with very different climate sensitivities can nevertheless produce the same centennial trend in air temperature.

   Any look at a perturbed physics study, such as [*Rowlands et al.*, 2012; *Yamazaki et al.*, 2013] show large disparities among climate models and model runs.

Therefore the projection variability to which the reviewer alludes is present, but is suppressed by model tuning, and is unappraised by the way that projection uncertainty is assessed in the field, namely as relative to an ensemble mean rather than as relative to physical accuracy.

These disparities among models are due to parameter uncertainties, only. They do not reveal errors due to incorrect or incomplete physical theory. Nor do they reveal the disparities relative to a physically true climate state.

The manuscript analysis shows that the annual average ±4 Wm^{-2} LWCF uncertainty arises within the models themselves. It is not part of F_0 or dF_i. As an inherent model error, LWCF uncertainty enters into every single step in a simulation, and necessarily propagates into and through a projection. The inevitable result is an expanding uncertainty envelope in a climate futures projection.

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.

Helton, J. C., J. D. Johnson, W. L. Oberkampf, and C. J. Sallaberry (2010), Representation of analysis results involving aleatory and epistemic uncertainty, Int'l J. Gen. Sys., 39(6), 605-646, doi: 10.1080/03081079.2010.486664.

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.

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.

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.

Roy, C. J., and W. L. Oberkampf (2011), A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing, Comput. Methods Appl. Mech. Engineer., 200(25-28), 2131-2144, doi: 10.1016/j.cma.2011.03.016.

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.

Taylor, B. N., and C. E. Kuyatt. (1994), Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, *NIST Technical Note 1297 Rep.*, 20 pp, National Institute of Standards and Technology, Washington, DC.

Vasquez, V. R., and W. B. Whiting (1998), Uncertainty of predicted process performance due to variations in thermodynamics model parameter estimation from different experimental data sets, Fluid Phase Equilib., 142(1-2), 115-130.

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.

Yamazaki, K., et al. (2013), Obtaining diverse behaviors in a climate model without the use of flux adjustments, J. Geophs. Res. - Atmos., 118(7), 2781-2793, doi: 10.1002/jgrd.50304.