Patrick Frank 28 February 2016
Propagation of Error and the Reliability of Global Air Temperature Projections
Manuscript # JCLI-D-15-0797

Author Response to Reviewer 2.

Summary

The reviewer:

- has completely, utterly, and invariably missed the critically central point that linear extrapolation of forcing directly entails linear propagation of error;
- neither understands nor respects the critically important distinction between accuracy and precision (items 1.1; 2.1.1-2.1.5; and 3.1-3.9);
- does not understand propagation of error (items 2.1.1, 2.1.6., and 2.2.2.2);
- does not understand the critical difference between error and uncertainty (items 2.2.1., 2.2.2.1, and 3.1-3.9);
- does not understand that compensating errors do not remove projection uncertainty (items 4.2 and 4.3);
- has implicitly conceded the validity of the author's analysis (item 5.8);
- has misconstrued emulation of models with climate physics, mistakes a statistic for an energy term, and supposes a random walk can proceed in two opposed directions simultaneously (items 6.1-6.4).

Detailed Response:

The reviewer is quoted in full in italics, followed by the author response. Review paragraphs, and sometimes sentences, are divided to attend individual points.

Prologue*. This manuscript is an impostor. It pretends to be a research paper, but it is not. A research paper makes a good-faith effort to provide relevant context as defined by prior work and then goes on to show how the prior work is extended (perhaps by showing flaws in it). By contrast, the present manuscript uses scientific-looking reasoning and referencing but blatantly ignores or misrepresents prior work. Furthermore, the manuscript contains fundamental errors in the technical development it claims to present as novel. Hence I cannot but recommend rejection.*

> The reviewer's opening comments are defamatory and of a construction to prejudice the editor. As shown below, the manuscript neither ignores nor misrepresents prior work. Nor does it contain fundamental errors. In view of the summary above, and the assessment below, the reviewer's prologue constitutes an inadvertent irony.

1. *The framing is disingenuous.*

> 1. The opening comment is prejudicial and violates ethics by an unjustifiable libel.

1.1. Following a superfluous and lengthy reminder of the distinction between precision and accuracy,...

1.1. The distinction between accuracy and precision is fundamental to science. Including the definition is hardly superfluous when throughout the published corpus of climate modeling, precision is conflated with accuracy.

Everywhere, variance about a model ensemble mean is represented as predictive uncertainty. However, these precision magnitudes convey nothing of accuracy and are therefore physically meaningless.

When such incorrect representations are nearly universal in the field, attention to the distinction is critical, not superfluous.

1.1.2. As shown below, reviewer 2 consistently confuses precision with accuracy; evincing no understanding of the distinction.

1.1.3. The accuracy vs. precision description consists of but one paragraph; this is hardly "*lengthy*."

1.2. *... it is first claimed that error propagation is ignored in the discussion of uncertainty of climate projections (line 110 onward).*

1.2. Lines 110-115 demonstrate the lack of propagated error by reference to published studies. The reviewer has ignored this evidence.

Error propagation is described analytically in manuscript eqns. 1&2, the like of which never appear in assessments of climate model projections.

1.3. *Then the quote from Smith (2002) is used to insinuate that climate modelling ignores elementary (high-school level) codes of good scientific practice concerning uncertainty estimates (line 117 onward).*

1.3. Nothing is insinuated. Line 117ff correctly notes that climate model assessments universally represent precision as predictive uncertainty. However, uncertainty bars constrained to precision do not reflect accuracy. They do not communicate predictive reliability.

Propagation of physical error is universally absent from assessments of climate models. Noting this fact is not an insinuation.

1.4. *Both statements are so misleading that they are effectively wrong (see major comments 2 and 3).*

1.4. As described above, and as shown below, manuscript lines 117ff are neither misleading nor wrong.

2.1. *Linear propagation of information and, in principle, error is investigated systematically by the so-called tangent-linear model or its transpose, the adjoint (e.g., Hall et al. 1982; Errico 1997; Marotzke et al. 1999).*

2.1. Linear propagation of information (the adjoint response) is not propagation of error, even in principle.

2.1.2. The reviewer's evidentiary citations are considered in turn below. The reviewer will be shown to have conflated precision with physical error; a mistake that the discussion in lines 72-81 of the Introduction were meant to forestall. These lines, the reviewer judged superfluous.

2.1.3 (Hall et al. 1982): These authors describe the adjoint model approach as a sensitivity analysis, even in the title of their paper.

They note (p. 2038) that, "*sensitivities quantify the extent that uncertainties in parameters contribute to uncertainties in results of models.*"

Further (p. 2039), "*The [adjoint] method starts by expressing a result, customarily called the response, as a functional of the model variables. ... The sensitivity of the response to all the model parameters can be evaluated ...*"

I.e., adjoint model parameter sensitivity tests determine the extent of model variance, a measure of precision. They do not indicate physical error with respect to the climate itself.

2.1.4. (Errico 1997): These authors describe the use of adjoint models as (p. 2578), "*The principal application of adjoint models is sensitivity analysis, and all its other applications may be considered as derived from it.*"

Once again, adjoint sensitivity analysis is said to assess model variance; a measure of precision, not error.

2.1.5. Likewise, (Marotzke et al. 1999) p. 29,530 column 1, "*An adjoint tracks the sensitivity of output with respect to input ...*" again constraining the meaning of an adjoint model to precision.

2.1.6. The transformation of climate model equations into their adjoint reveals model sensitivities. It does not propagate error through projections of climate futures.

2.1.7. Nor is the adjoint transformation equivalent to the manuscript demonstration that models project temperature by linear extrapolation of greenhouse gas forcing.

2.1.8. Manuscript line 64 points out that sensitivity analyses reflect model precision, not model accuracy; nor are they propagated error. The reviewer ignored this.

2.1.9. The reviewer has evinced no understanding of the critical distinction of accuracy from precision, making the reviewer's entire line of reasoning in item 2.1 misguided and an irrelevant digression.

2.2.1. *However, application to a full-blown climate model is conceptually and technically extremely challenging, because turbulent instabilities limit the utility of the linearization (e.g., Lea et al. 2000; Köhl and Willebrand 2002) and because accumulation of systematic error is very hard to trace quantitatively (e.g., Rauser et al. 2011).*

2.2.1. The manuscript is about the accumulation of uncertainty, not about the accumulation of error. The two are not the same at all; a critical point the reviewer has missed.

Nor evidently has the reviewer consulted Section 10 in the Supplemental Material, where the meaning of uncertainty is discussed in detail.

2.1.2. Model results are proved to be a linear extrapolation. Quantifying the growth of uncertainty by linear propagation of error is completely standard practice.

2.1.3. Advanced climate models project temperature by linear extrapolation of forcing. This is thoroughly and definitively demonstrated in manuscript Section 2.2, Figure 2, Figure 3, and Figure 4, and Section 2.4.2 Figure 8, along with Supplementary Material Figure S3, Figure S4, Figure S5, Figure S6, Figure S7, and Figure S8.

Projection output remains linear across models despite variations in parameter values and despite "*turbulent instabilities*."

The reviewer has completely ignored this demonstration. The reviewer has also ignored that linearly cumulated outputs directly entail linear propagation of error.

The reviewer's comments about turbulent instabilities are entirely irrelevant to the demonstration of model linear temperature projections and to linear propagation of error.

2.2.2 *The premise stated here, that linear error propagation is ignored in climate modelling, is hence plain wrong.*

2.2.2.1 The significance of manuscript eqns. 1 and 2 has clearly escaped the reviewer, as has the distinction between uncertainty and error.

The reviewer's mistaken equation of adjoint sensitivity analysis with linear propagation of physical error is fatal to objection 2.

Manuscript eqn. 2 provides the analytical form of propagated error. It is not used in a sensitivity analysis. The reviewer has failed to show any equivalence; such an equivalence cannot be shown even in principle

2.2.2.2. Indeed adjoint sensitivity analysis is not even propagation of precision variance.

For example, suppose an adjoint analysis produced a suite of variations of individual model runs around an ensemble mean. For this purpose, we call precision "error" and is defined as $(y - \mu)_i$, i.e., the difference between each output and the corresponding ensemble mean value calculated at each i^{th} time-step for each run of *n* steps.

Manuscript eqn. 2 tells us that the uncertainty due to lack of precision at each i^{th} time-step for each model run is propagated as, $\pm\sigma_u^p = \sqrt{\sum_{i=1}^{n}(\mu - y)_i^2}$, which would appear as serially increasing uncertainty bars around the model trend.

Applied to the ensemble mean itself, the total precision uncertainty at each i^{th} time-step becomes the rms of the individual model uncertainties at each step *i*,

$\pm\sigma_{\mu_i}^p = \sqrt{\frac{\sum_{j=1}^{m}(\pm\sigma_{u_i}^p)_j^2}{m}}$, where *j* indexes the individual model run across *m* runs and $\pm\sigma_{\mu_i}^p$ serially increases in width along the mean trend.

However, this form of precision analysis is never reported in assessments of climate models. Instead, projection uncertainty is merely the progression of $(y - \mu)_i$ about the ensemble mean.

Thus the reviewer has both mistakenly equated sensitivity studies to propagation of physical error and mistakenly supposed that the precision width from a sensitivity study represents propagated uncertainty.

This double mistake evidences a reviewer in complete ignorance of error propagation.

2.2.2.3. One observes that propagating precision in the absence of an accuracy metric is a physically meaningless irrelevance, and in any case is something that is never done in climate model studies.

2.2.3. *What the author would need to show is that he can overcome the formidable limitations identified to date.*

2.2.3. The author showed exactly that.

As evidence, the reviewer is referred to manuscript Section 2.2, Figure 2, Figure 3, and Figure 4, and Section 2.4.2 Figure 8, along with Supplementary Material Figure S3, Figure S4, Figure S5, Figure S6, Figure S7, and Figure S8, followed by the linear propagation of long-wave cloud forcing error displayed in Section 2.4.2 Figure 7 and Figure 8 and Section 2.4.4 Figure 9.

Climate models are demonstrated to project temperature as a linear extrapolation of greenhouse gas forcing. Linear propagation of error follows immediately.

The above logic should be directly evident to any working physical scientist.

However, the reviewer evidently did not grasp that obvious demonstration.

3. *Any serious look, however brief, into the IPCC report shows how carefully uncertainty is estimated, given the inherent limitations. For example, Figure 12.1 of WGI AR5 Chapter 12 (Collins et al. 2013) lists the ensembles of the state-of-the-art climate models, and the surrounding text goes to great lengths in explaining how uncertainty of the projections is estimated. Ignoring this, as the present manuscript does, is disingenuous.*

3.1. Collins, et al., (2013) AR5 WG1 Chapter 12 discusses uncertainty in terms of model variability; that is, model precision (cf. 3.3ff).

The manuscript evaluates model accuracy. Model precision has no relevance whatever to this work. Therefore, ignoring precision is not disingenuous, but is by necessity.

3.2. Reviewer comment 3 centers on precision; the manuscript centers on accuracy. This stark disparity evinces the reviewer does not grasp the absolutely fundamental distinction between accuracy and precision.

3.3. Examples from Ar5 WG1 Chapter 12: Section 12.2.2 (p. 1039) discusses uncertainty in terms of model spread around an ensemble mean ("*model range*" or "*model mean*") and with respect to model representation of the physical climate ("*model uncertainty*").

Section 12.2.2 goes on to say, "*For forecasts of global temperatures after mid-century, scenario and model ranges dominate the amount of variation due to internally generated variability, with scenarios accounting for the largest source of uncertainty in projections by the end of the century.*"

This "*model range*" uncertainty is variance about an ensemble mean, i.e., the AR5 WG1 uncertainty here is the $\pm(y - \mu)_i$ model precision noted above, not physical error and certainly not propagated model error.

"*Model uncertainty*" can be a physical error metric as discussed in manuscript lines

62-70. However, physical error is never propagated through model projections and never makes an appearance in the uncertainty bars about projected temperatures.

3.4. More AR5 precision: WG1 Figure 12.5 provides air temperature projections and their uncertainties. From Figure 12.5 Legend:

"*Projections are shown for each RCP for the multi-model mean (solid lines) and the 5 to 95% range (±1.64 standard deviation) across the distribution of individual models (shading).*"

These model standard deviations are described as projection uncertainties in Table 12.2. They represent model precision, not accuracy.

3.5. More AR5 precision: Figure 12.8 shows a set of uncertainties in projected 2081-2100 air temperature, calculated by a variety of methods. Every single method represents a model spread. I.e., model precision.

3.6. More AR5 precision: WG1 Annex 1"*Atlas of Global and Regional Climate Projections*" p. 1313 says,

"*Here the range of model spread is provided as a simple, albeit imperfect, guide to the range of possible futures (including the effect of natural variability). Alternative approaches used to estimate projection uncertainty are discussed in Sections 11.3.1 and 12.2.2 to 12.2.3.*"

Model spread is a precision metric. That is, all the methods of determining uncertainty involve precision, as noted above.

3.7. More AR5 precision: WG1 Section 11.3 is "*Near Term Projections*" which projections extend to 2100. "*Section 11.3.1.1 Uncertainty in Near-term Climate Projections*"

Therein, Figure 11.8 provides projection uncertainties which are described as arising, "*from internal variability (orange), model spread (blue) and RCP scenario spread (green).*" I.e., the uncertainties represent model spread about a mean: precision. They do not represent physical error or propagated error.

3.8. More AR5 precision: Further uncertainties are discussed in AR5 WG1 13.3.6.1, quantified in section 13.3.6.3 and displayed in Figure 11.25. Every single one of them involves model spread: precision.

3.9. Error propagation as given in manuscript eqn. 1 and eqn. 2, and as employed throughout the work, appears nowhere in AR5 WG1 Chapter, 10, 11, or 12.

4. *Claiming that the well-known systematic error in cloud radiative effect (CRE) swamps any smaller energy flux ignores the presence of compensating errors (e.g., WGI AR5 Ch9,*

page 766; Flato et al. 2013).

4.1. The only mention of compensating errors in AR5 WG1 Ch 9, p. 766 is:

> "*Regime-oriented approaches to the evaluation of model clouds (see Section 9.2.1) have identified that compensating errors in the CRE are largely a result of misrepresentations of the frequency of occurrence of key observed cloud regimes...*"

Nowhere in Flato, et al. 2013 (AR5 WG1 Chapter 9) where compensating errors are mentioned (pp. 761, 766, 785) are they said to remove projection uncertainty.

4.2. The reviewer does not seem to realize that the reproduced observables owing to compensating errors to do not imply that the climate state is simulated correctly, nor that the physics is correct, and certainly do not imply that the climate will be projected correctly.

When the physics is incorrect, quantifiable uncertainty still remains in a hindcasted observable no matter that it has the correct magnitude, because the physical description of the underlying climate energy state is not known to be correct. The statistical merit of the observable therefore does not indicate a state of knowledge.

4.3. The reviewer here has implicitly raised the issue of model tuning. Tuned parameter sets amount to offsetting errors, and are the reason observables can be reproduced in a hindcast. Model tuning does not at all reduce or remove uncertainty, because no errors in the model physics were corrected.

The problem is discussed in detail in Supplementary Material Section *7.3 The problem of tuning and parameter uncertainty*. However, this discussion apparently escaped the reviewer's notice (despite a Table of Contents).

The heading of manuscript Section 2.4.3 has been modified and a sentence has been added there to draw attention to SM Section 7.3 and the problem that model tuning and compensating errors do not remove projection uncertainty.

5. *Section 2.1 goes through a tortuous derivation of the decades-old linear relationship between forcing and response (e.g., Gregory et al. 2004) - but arriving at a fundamentally wrong version of this relationship (see below).*

5.1. Section 2.1 derives the onset of significant CO_2 forcing. An extensive check of textbooks and of prior literature nowhere revealed a similar evaluation. Therefore, the derivation is not decades old, but original with this manuscript.

5.2. The same check revealed that nowhere else is CO_2 forcing discussed in terms of the 15µ IR photonic mean free path.

5.3. Eqn. 6 was derived from the CO_2-forcing relationship expressed within a climate model, following a completely different logic than that of Gregory, et al. 2004.

5.4. (Gregory et al. 2004) refer to their equation as reproducing "*any GCM experiment in which the climate is responding to a constant forcing*. (underline added)"

In contrast, manuscript equation 6 emulates any GCM simulation with a varying forcing, or indeed any forcing at all.

Figure S11, for example, emulates the a GCM projection that includes forcing changes due to volcanic aerosols. Figure S12 and Figure S13 emulate the entire 20th century record using the varying IPCC forcings (compare Figure S14, which shows GCM simulations of the same record).

5.5. Demonstrating the difference strictly: the Gregory, et al. 2004 equation is $N = F - \alpha\Delta T$, where N is the net downward heat-flux in Wm^{-2}, α is "*the climate response parameter,*" a proportionality constant in units of $Wm^{-2}K^{-1}$, F is CO_2 forcing in Wm^{-2}, and ΔT the change in air temperature due to N, in Kelvins.

Rearranging, $\Delta T(K) = (F-N)/\alpha = \alpha'\times(F-N),$ where $\alpha' = 1/\alpha$, in $W^{-1}m^{2}K$.

Compare manuscript eqn. 6, condensed form: $\Delta T(K) = a(K)\times(\Delta F/F_0)$. The proportionality constant *a,* in Kelvins, operates on the fractional change in forcing, which is dimensionless.

The differences include the functional form, the terms, and the internal dimensions.

Manuscript eqn. 6 clearly is not a version of the Gregory, et al., 2004 equation. The reviewer is mistaken.

As eqn. 6 is not a version of the Gregory relation, it categorically cannot be a wrong version of the Gregory relation.

5.6. Manuscript eqn. 6 is demonstrated fully able to successfully emulate the temperature projections made using advanced climate models. Therefore, it is not "*wrong*" at all.

5.7. Manuscript eqn. 6 emulates the air temperature projections of any advanced climate model, a capacity that was not demonstrated for the Gregory equation.

5.8. In comment 5, the reviewer has implicitly acceded to the validity of the linear propagation of climate model error, by resting an argument on the fact that Gregory, et al., showed GCM temperature projections can be approximated using a linear equation.

That is, once GCM outputs are known to be linear, then linear propagation of error is immediately justified. The reviewer has accepted the former, and is thus logically compelled to accept the latter.

6.1. *Equation (8), formulating an energy-balance framework leading to the relationship between forcing and response, leaves out a crucial term - the damping from changed radiation to space.*

6.1.1. Equation 8 does not formulate "*an energy-balance framework.*" Equation 8 merely generalizes eqn. 6. And eqn. 6 merely demonstrates the linearity of climate model output.

That is all it does, and all it was ever meant to do.

Eqn. 8 does not formulate any physics of climate. Nor does eqn. 6.

Lines 154-157 of the Introduction could not make this more plain:

> "*[eqn. 6] is not a climate model. The physics of climate is neither surveyed nor addressed ... the focus is on the behavior and reliability of climate models themselves.*"

This restriction is also plainly stated in manuscript lines 130-131, where the derived CO_2 forcing fraction is related to climate models, not to the climate itself.

The reviewer has overlooked these plain statements. The objection is misconceived and irrelevant.

6.1.2. The CMIP5 models presumably included, "*the damping from changed radiation to space*." Nevertheless, eqn. 6 successfully and invariably emulated their temperature projections. Where, then, is the further need for damping?

Note that despite the damping, CMIP5 models produced temperature projections linear with forcing. It thus appears that damping is linear with forcing.

6.1.3. Eqn. 8 and eqn. 6 do one thing only: emulate the observed linear output of climate models.

A successful emulation obviously does not require inclusion of any further term.

6.2. *Furthermore, there is no justification for adding a systematic error as a random "forcing" in a Markov process, because it is then implied that the errors are uncorrelated in time.*

6.2.1. How can a forcing can be simultaneously positive and negative, i.e., (plus/minus)4 Wm^{-2}? The reviewer's diagnosis clearly violates physics; ±4 Wm^{-2} cannot be a forcing.

6.2.2. Temperature projection reflects a deterministic process, not a Markov process.

6.2.3. The ±4 Wm^{-2} longwave cloud forcing (LCF) error ("*systematic error*") is a statistical time-average of 520 simulation years (Lauer and Hamilton 2013) from 26 CMIP5 GCMs and is therefore time-independent.

LCF error is present in every single CMIP5 simulation time-step. As an annual average LCF error statistic, ±4 Wm^{-2} is representative of the uncertainty in any CMIP5 simulation.

As a representative annual model uncertainty, it is entirely justifiable to propagate ±4 Wm^{-2} through any CMIP5 simulation in order to determine a representative rate of annual growth of uncertainty.

6.2.4. The reviewer has wrongly characterized the ±4 Wm^{-2} LCF uncertainty. It is not a random forcing, or any sort of forcing. It is not an energy term at all.

The ±4 Wm^{-2} is a statistic. It does not influence the model or enter into the simulation.

6.2.5. Propagation of an average uncertainty statistic does not imply anything about correlation of error in time. It does not imply anything at all about error. Propagation converts model uncertainty into projection uncertainty.

Obviously, nothing can be known of projection error magnitude or correlation because in a futures simulation the error is not knowable.

Manuscript line 566 notes that growth of uncertainty conveys an increasing level of ignorance about the validity of the simulated climate state. It does not represent a growth in error.

Uncertainty implies ignorance about outcomes. It indicates the growth of ignorance concerning the conformance of the phase-space position of the simulated future climate to that of the future physically real climate. Such ignorance necessarily increases with projection time.

This meaning of uncertainty is discussed in lines 692-714; however to no avail, apparently.

Manuscript line 683 explicitly recommends the extensive discussion of the meaning of uncertainty in Supplemental Material Section 10.2 *The meaning of predictive uncertainty*. Perhaps the reviewer overlooked this recommendation.

The manuscript has been revised to make the recommendation to this SM section more obvious.

6.2.6. Understanding the distinction between uncertainty and error is critical to the manuscript. However, the distinction is apparently unknown to reviewer.

6.3. *This latter property, together with the absence of damping, means that equation (8) simulates a simple random walk (e.g., Wunsch 1992), ...*

6.3.1. The reviewer may wish to contemplate how a random walk can simultaneously proceed in opposed directions. I.e., the (not)forcing is plus/minus 4 Wm^{-2}.

6.3.2. As noted in 6.2.3, the ±4 Wm^{-2} LCF error is a statistical time-average, not a random forcing and not an energy term. It therefore does not have the, "*latter property*" of physical error itself, much less one uncorrelated in time.

6.3.3. Likewise as noted in 6.1.2 and 6.1.3, "*the absence of damping*" is entirely irrelevant to the emulation function of eqn. 6 and eqn. 8. The equations do not express climate physics. They express, and successfully so, the observed output behavior of climate models.

6.4 *... and its standard deviation after 100 years comes out trivially for the current example as 0.42 x 4 x (number of years)^(1/2) K ca. 16 K. Thus, the manuscript with its fatally flawed energy balance merely repeats a decades-old example from random walk.*

6.4.1. Correcting the reviewer's example following from 6.3.1: $0.42 \times \pm 4 \times (100)^{1/2}$ K ≈ ±17 K. The plus/minus result is not a random walk.

6.4.2. The reviewer's equation omits the greenhouse temperature and forcing terms of eqn. 6. Therein, the 33 (K) of greenhouse warming is nearly normalized by the 30.45 (Wm^{-2}) of year 1900 forcing. This happens to make the annual increase in uncertainty nearly constant. The reviewer's result is thus fortuitous.

6.4.3. The reviewer's evident inability to distinguish between an uncertainty statistic and a forcing is fatal to the reviewer's argument, and indeed to the entire review.

References

Errico, R. M., 1997: What Is an Adjoint Model? *Bulletin of the American Meteorological Society*, **78,** 2577-2591.

Gregory, J. M., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31,** n/a-n/a.

Hall, M. C. G., D. G. Cacuci, and M. E. Schlesinger, 1982: Sensitivity Analysis of a Radiative-Convective Model by the Adjoint Method. *Journal of the Atmospheric Sciences*, **39,** 2038-2050.

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,** 3823-3845.

Marotzke, J., R. Giering, K. Q. Zhang, D. Stammer, C. Hill, and T. Lee, 1999: Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *Journal of Geophysical Research: Oceans*, **104,** 29529-29547.