**Keeping it simple: the value of an irreducibly simple climate model**

**Christopher Monckton of Brenchley · Willie W-H Soon · David R Legates · William M Briggs**

**Abstract** Richardson *et al.* (2015) suggest that the irreducibly simple climate model in Monckton of Brenchley *et al.* (2015) was not validated against observations, relying instead upon synthetic test data based on subjective assumptions including an underestimation of warming, an illogical parameter choice and a near-instantaneous response to forcing at odds with ocean warming and other observations. However, the simple model informed by our choice of parameters performs better against recent observed temperature change than the general-circulation models. Its predictions are closer to observations than theirs. When informed by the IPCC’s choice of parameters, the simple model is indeed validated in that, as Richardson *et al.* concede, it correctly replicates the IPCC’s sensitivity interval. Furthermore, a near-instantaneous response to forcing is an ineluctable consequence of a near-zero or net-negative temperature feedback sum. Given the large uncertainties in both the initial conditions and the time-dependent evolutionary processes determinative of climate sensitivity, subject to obvious caveats a simple sensitivity-focused model is not inherently likely to exhibit significantly less predictive skill than the general-circulation models. The simple model finds high climate sensitivity estimates implausible.

**Keywords** Climate change · climate sensitivity · climate models · global warming · temperature feedbacks

1. **Introduction**

**1.1 Defects in complex models’ output** Outputs from the general-circulation models cited in IPCC’s five *Assessment Reports* FAR, SAR, TAR, AR4 and AR5 [1-5] were examined in [6] using a simple model calibrated against IPCC’s central climate-sensitivity estimates in AR4-5 and were found to overestimate observed air temperature trends. Reasons for this hot running were discussed. The simple model, using parameter values consistent with a growing body of papers that report temperature feedbacks to be net-negative, showed that in at least five significant respects the general-circulation models’ approach was questionable:

* The assumption that temperature feedbacks will double or triple direct warming is unsound. Feedbacks may well reduce warming, not amplify it (see e.g. [7-8]).
* The Bode system-gain equation models mutual amplification of feedbacks in electronic circuits, but, when models erroneously apply it to the climate on the assumption of strongly net-amplifying feedbacks, its use leads to substantial overestimation of global warming (see [9] for a discussion).
* Climate modelers have failed to cut their central estimate of global warming in line with a new, lower estimate of the feedback sum (AR5, fig. 9.43). They still predict 3.3 K warming per CO2 doubling, when on this ground alone they should only predict 2.2 K, about half from direct warming and half from feedbacks.
* Though general-circulation models suggest 0.6 K manmade warming is “in the pipeline” even if CO2 emissions cease, the simple model, supported (see Fig. 7) by almost two decades without significant global warming [10], suggests there is no committed but unrealized manmade warming still to come.
* AR5’s extreme RCP 8.5 forcing scenario predicting ~12 K anthropogenic warmingis unjustifiable. It was based on CO2 concentration growing at 5.5 ppmv yr–1 this century, though AR4’s central estimate was below 3.5 ppmv yr–1 and the current growth rate [11] is little more than 2 ppmv yr–1.

In [6], it was concluded that once due allowance is made for these and other shortcomings in the general-circulation models, the likely global warming in response to a doubling of CO2 concentration is not 3.3 K but 1 K or less, and that even if all available fossil fuels are combusted less than 2.2 K warming will result.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

C Monckton of Brenchley ()

Science and Public Policy Institute, 5501 Merchants’ View Square #209, Haymarket, VA 20169, United States of America

e-mail: [monckton@mail.com](mailto:monckton@mail.com). Telephone +44 7814 556423

W Soon

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, United States of America

(for identification only: the co-author’s opinions are his own)

D Legates

Department of Geography, University of Delaware, Newark, DE 19716, United States of America

W Briggs

New York, NY 10021, United States of America

**1.2 The simple model** The irreducibly simple model presented in [6], encapsulated in (1), determines the surface temperature response Δ*Tt* to anthropogenic radiative forcings and consequent temperature feedbacks over any given period of years *t*:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  | (1) |
|  |  |  |  |
|  |  |  |  |

whereis the fraction of total anthropogenic forcing represented by CO2 over *t* years, and its reciprocal allows for non-CO2 forcings as well as the CO2 forcing; Δ*Ft* is the radiative forcing in response to a change in atmospheric CO2 concentration over *t* years, which is the product of a constant *k* and the proportionate change (*Ct* / *C*0) in CO2 concentration over the period ([12] ; TAR, ch. 6.1); *rt* is the transience fraction, which is the fraction of equilibrium sensitivity expected to be attained over *t* years; and *λ*∞ is the equilibrium climate-sensitivity parameter, which is the product of the Planck or zero-feedback sensitivity parameter *λ*0 (AR4, p. 631 fn.) and the open-loop or system gain *G,* which is itself the reciprocal of 1 minus the closed-loop gain *g,* which is in turn the product of *λ*0 and the sum *ft* of all temperature feedbacks acting over the period.

In [6] the simple model encapsulated in (1) is described thus: “This simple equation represents, in an elementary but revealing fashion, the essential determinants of the temperature response to any anthropogenic radiative perturbation of the climate, and permits even the non-specialist to generate respectable approximate estimates of temperature response over time. It is not, of course, intended to replace the far more complex general-circulation models: rather, it is intended to illuminate them.”

**2 Criticisms of the simple model**

In [13], various criticisms of the simple model in [6] were tendered. We now consider those criticisms *seriatim*.

**2.1 Form of the simple model** It is stated in [13] that the simple model’s “extreme simplification necessarily leaves out many physical processes”. The model was intentionally simple. The aim was to allow an undergraduate student of climatological physics to understand the key forcings, feedbacks and other parameters determinative of climate sensitivity and to generate respectable climate-sensitivity estimates that would serve to illuminate the outputs of the general-circulation models. General circulation models are inherently complex, so that processes and feedbacks are often obscured. The goal of the simple model is to be transparent and to allow for direct evaluation of the response function associated with its components. Indeed, so large are the uncertainties in the initial conditions and in the time-dependent processes by which the climate object evolves that the penalty in predictive skill arising from simplicity is small. For every model is a simplification, and every simplification is an analogy, and every analogy breaks down at some point. Naturally, [6] contained appropriate caveats about the limitations on the competence of a simple model.

**2.2 Focus on anthropogenic forcings** It is said in [13] that “the standard approach” that includes natural as well as anthropogenic forcings “is more useful” than the simple model “because both because both Δ*Tt* and Δ*Ft* may be estimated from observations”. However, reliable global temperature measurements only became available in 1979, during the satellite era. Uncertainties before that time do not provide a secure basis for analysis. Also, as [6] explained, the process of determining forcings and distinguishing them from feedbacks is subject to very large uncertainties. Even the usual central estimate of the CO2 forcing was reduced by 15% on the basis of intercomparison between three models in [12]. In any event, it is a simple matter for the user of the simple model to incorporate terms for natural forcings where that is thought desirable.

**2.3 The transience fraction** It is argued in [13] that the transience fraction *rt* should have been determined not by reference to a single pulse of forcing but by “a convolution of the forcing series with the temporal response function”. Naturally, one could have presented a more complex model. However, the effect of assuming a single pulse rather than a time-series of forcings is to increase the value of *rt* somewhat, for the sake of caution*,* where *t* is sufficiently distant both from zero and from equilibrium. At equilibrium, the primary focus of [6], *rt* is by definition unity. Likewise, if the feedback sum *ft* is below 0.1, no large error arises from assuming *rt* = 1. And it is made clear in [6] that the user is free to adopt any chosen time-series array of values for *rt*. Finally, it is ironic that the authors of [13] have failed to perform the calculation they had themselves recommended.

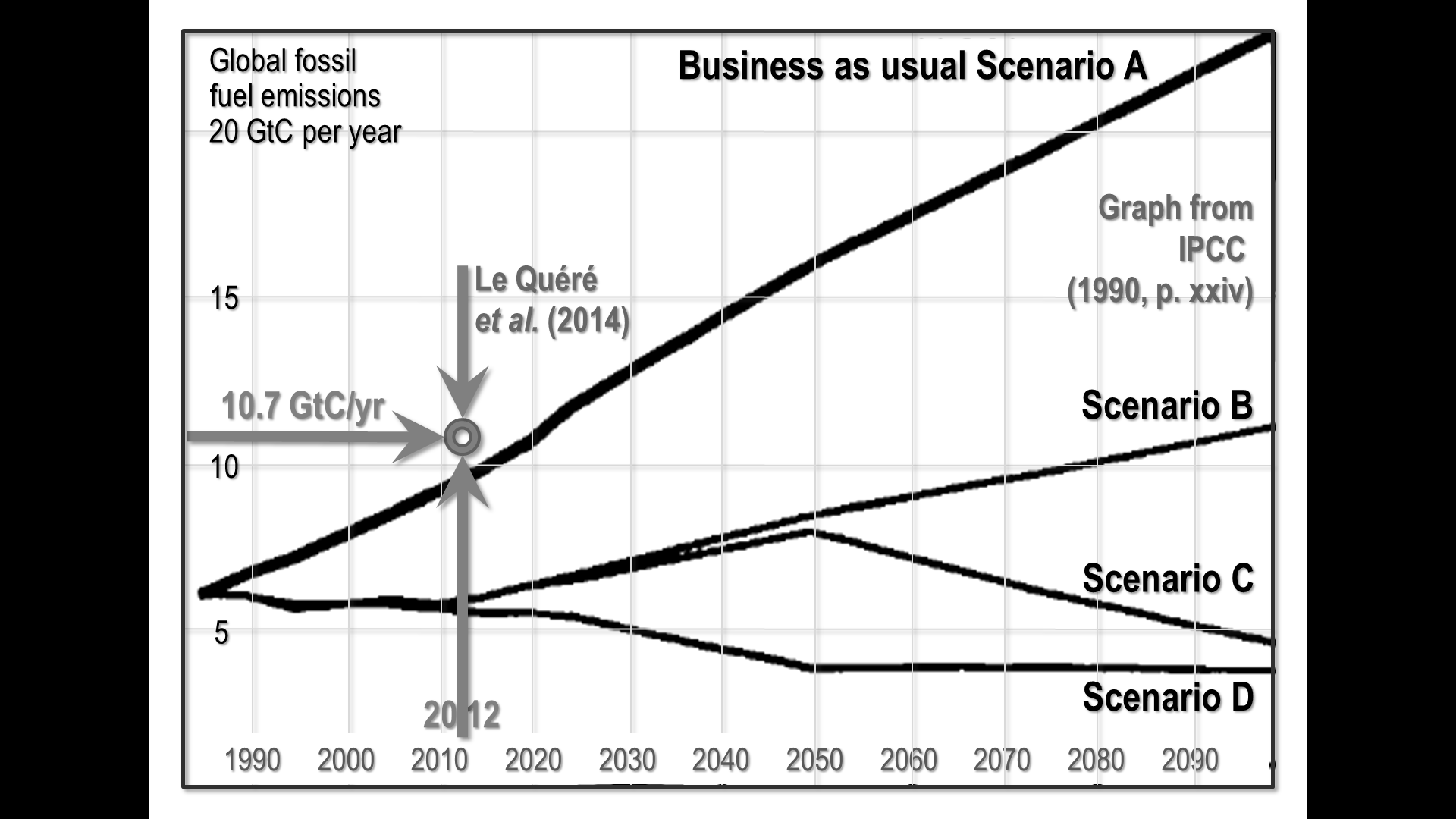
**2.4 Calibration of the simple model** It is argued in [13] that the projections of the simple model were not compared with observations. However, as explained in [6] and actually confirmed in [13], the model was calibrated against the central climate-sensitivity projections in AR4-5 and was found to perform satisfactorily. In [6] parameter values somewhat different from IPCC’s values were then selected, assuming net-negative feedbacks and also assuming a transience fraction *rt* = 1. The model was then run, determining a climate sensitivity of 1.0 [0.8, 1.3] K per CO2 doubling, in response to a CO2 radiative forcing 5.35 ln(2) = 3.7 W m–2.

AR5 (at Fig. SPM.6) estimates 2.3 W m–2 radiative forcings from all sources from 1750-2012. Over that period, some 0.9 K global warming has occurred, where the simple model in [6] would have led us to expect 0.6[0.5, 0.8] K. On this test, then, the simple model ran rather cool, though the upper bound of its predicted interval falls on the error interval of observed temperature change since 1750.

For comparison, the central sensitivity estimate derived from an ensemble of 18 models in AR4 (p. 798, box 10.2) is 3.3[2.5, 4] K. Accordingly, if 23/37 of that warming were to occur at equilibrium in response to 2.3 W m–2 of forcing, at a transience fraction *rt* = 0.4 (a suitable value on the high-sensitivity case, against *rt* = 1 in the low-sensitivity case), the general-circulation models would predict warming since 1750 on 0.8[0.6, 1.0] K.

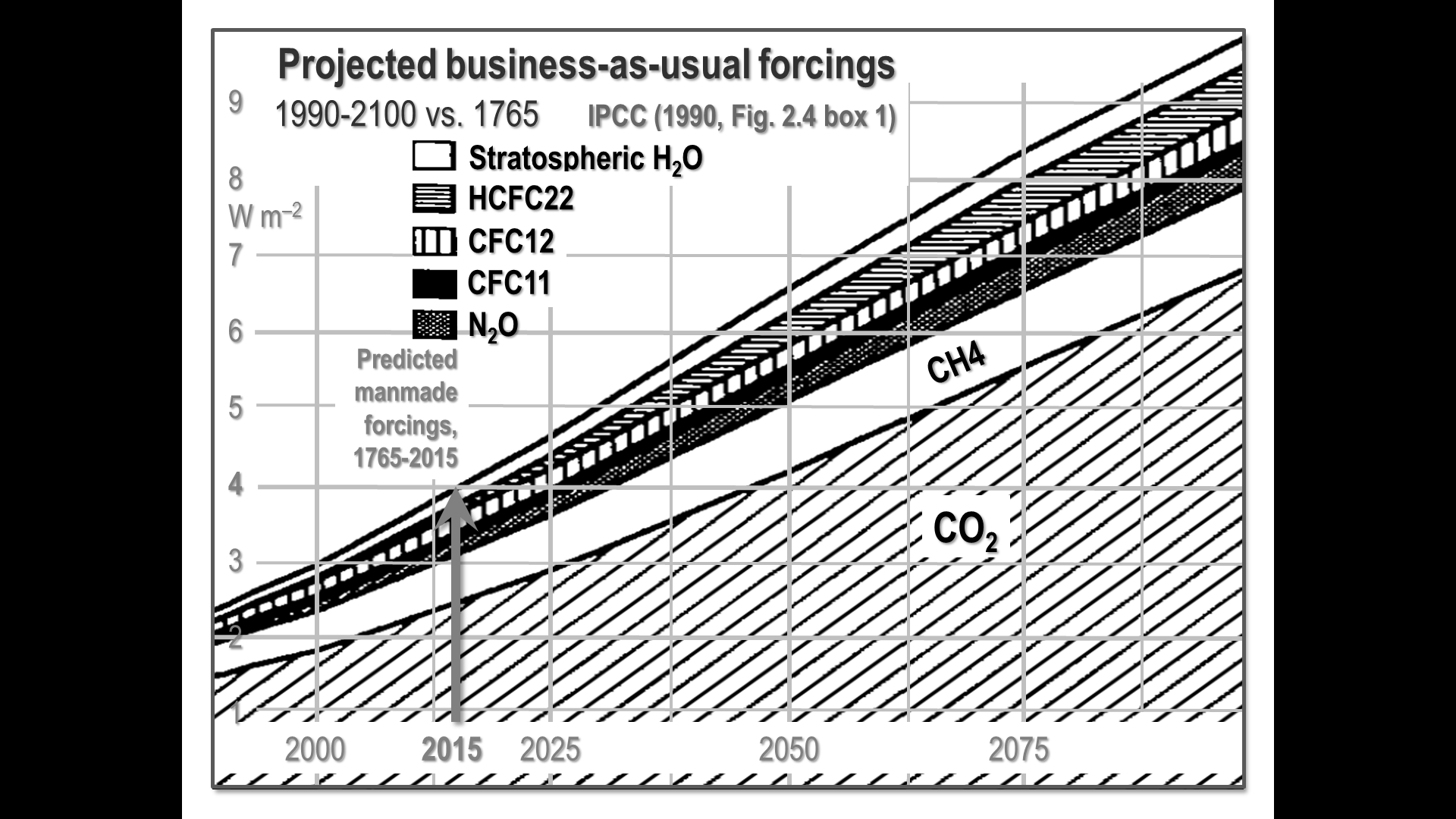
This admittedly back-of-an-envelope example, in which the transient-sensitivity intervals of the simple and general-circulation models are in fact substantially coincident, nicely illustrates the difficulties in empirically distinguishing the high-sensitivity and low-sensitivity cases at this stage. We further emphasize that we have set aside the still problematic tendency of the global temperature datasets to over-estimate the amplitude of anthropogenic warming over the past 150-250 years owing to contributions from non-climatic factors [14-15].

Worse, we do not really know – perhaps even to within a factor 2 – what is the magnitude of the total forcings since 1750. IPCC (1990, p. xxiv) provided a graph predicting future annual emissions of CO2 (Fig. 1). Scenario A, the “business-as-usual” case, predicted global emissions of almost 10 GtC yr–1 by 2012. In [16], it was estimated that in 2012 global emissions were actually 10.7 GtC, somewhat above even the business-as-usual prediction in FAR. That case, then, is closer than the other three FAR cases to the observed outturn.



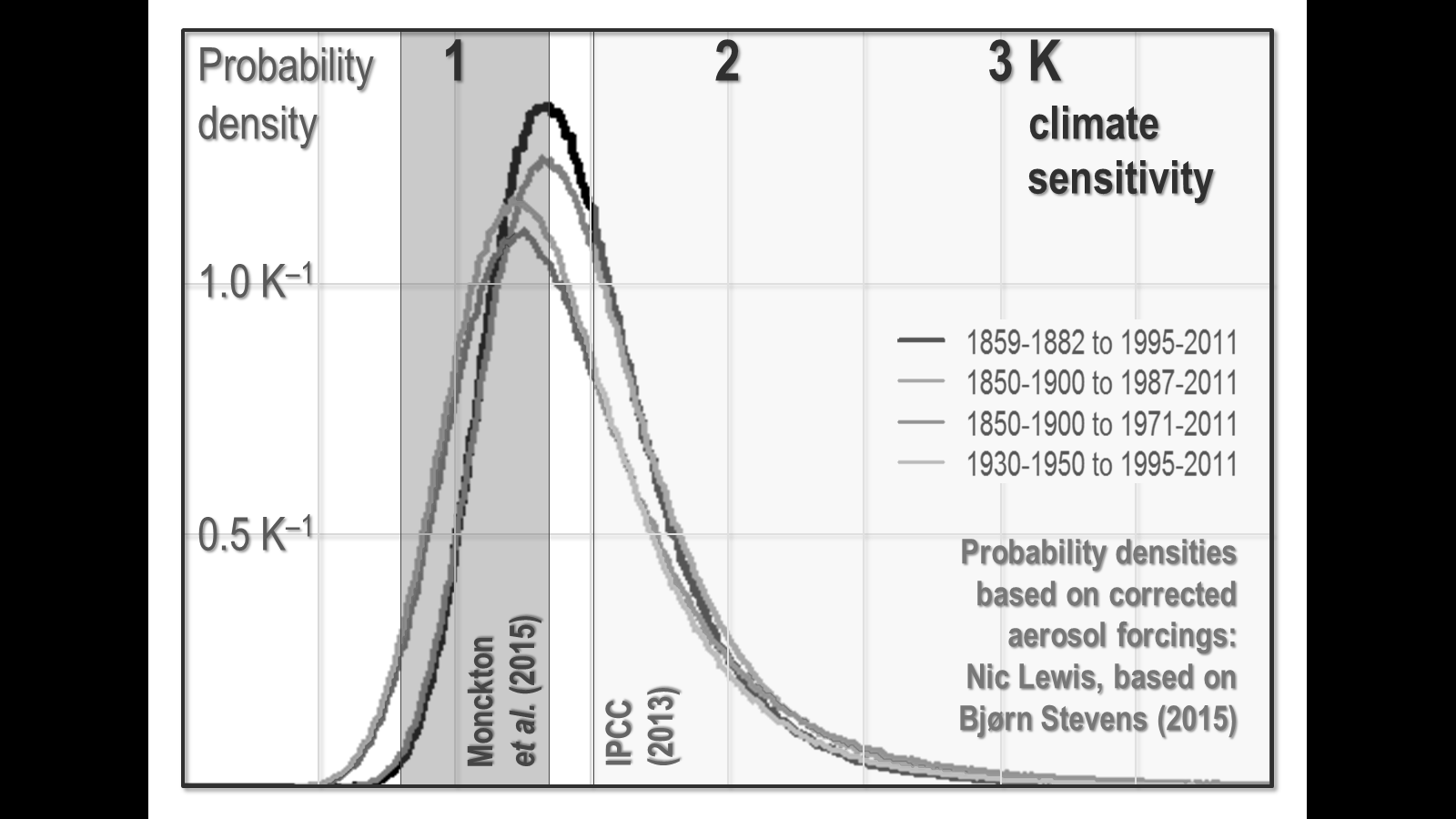
**Fig. 1** Predicted annual CO2 emissions (GtC) on four scenarios (FAR, p. *xxiv*), and 2012 outturn

FAR, on the business-as-usual case, predicted anthropogenic forcings of about 4 W m–2 between 1765 and 2015 (Fig. 2). If that value is correct, then the transient sensitivity intervals that the simple model and the CMIP3 general-circulation model ensemble relied upon in AR4 might predict would be 1.0 [0.9, 1.4] K and 1.4 [1.0. 1.7] K respectively. Again, the 0.9 K outturn falls close to the predicted sensitivity intervals both in the simple and in the general-circulation models, though this time the simple model is closer to observation.



**Fig. 2** Projected business-as-usual radiative forcings, 1990-2100 vs. 1765 (FAR, Fig. 2.4 box 1).

Why the difference between the forcings of 4 W m–2 predicted in FAR and the estimated outturn of little more than half of that value in AR5? The chief reason is that from SAR onward IPCC introduced the notion that anthropogenic aerosol forcings were pronouncedly negative, though the literature is divided on this subject. For instance, a recent paper [17] modeling climate sensitivity based on research showing that a strong aerosol forcing establishes itself early in the historical record finds sensitivity tightly constrained close to 1.3 K (Fig. 3).



**Fig. 3** Probability densities based on four different aerosol reanalyses in [18], indicating climate sensitivity is most likely to fall in the region of 1.3 K. Equilibrium climate sensitivity intervals in [6] and in AR5 are indicated by dark and light shadings respectively. Graph based on [19]

Given these and other uncertainties, one method of testing the simple model’s predictive skill is to compare its prediction of global warming in the 25 years 1990-2014 with that of the IPCC on the business-as-usual Scenario A, and also with observed outturn.

The executive summary of FAR asked, “How much confidence do we have in our predictions?” IPCC pointed out some uncertainties (clouds, oceans, etc.), but concluded:

“Nevertheless … we have substantial confidence that models can predict at least the broad-scale features of climate change … There are similarities between results from the coupled models using simple representations of the ocean and those using more sophisticated descriptions, and our understanding of such differences as do occur gives us some confidence in the results.”

FAR’s medium-term temperature-change prediction was as follows:

“Based on current models we predict:

“under the IPCC Business-as-Usual (Scenario A) emissions of greenhouse gases, a rate of increase of global mean temperature during the next century of about 0.3 Cº per decade (with an uncertainty range of 0.2 Cº to 0.5 Cº per decade), this is greater than that seen over the past 10,000 years. This will result in a likely increase in global mean temperature of about 1 Cº above the present value by 2025 and 3 Cº before the end of the next century. The rise will not be steady because of the influence of other factors” (p*. xii*).

Later, IPCC said:

“The numbers given below are based on high-resolution models, scaled to be consistent with our best estimate of global mean warming of 1.8 Cº by 2030. For values consistent with other estimates of global temperature rise, the numbers below should be reduced by 30% for the low estimate or increased by 50% for the high estimate” (p*. xxiv*).

Accordingly, IPCC predicted warming of 1.0 [0.7, 1.5] K to 2025 and warming of 1.8 [1.3, 2.7] K to 2030. Since FAR’s graph of the predicted warming is very nearly a straight line from 1990-2025, little error arises by assuming straight-line warming over the period, so that the pro-rata predicted warming interval from 1990-2014 was 0.7 [0.5, 1.0] K. However, observed warming over the quarter-century from 1990-2014, taken as the linear trend on the mean of the monthly anomalies on the three terrestrial and two lower-troposphere datasets, was 0.35 K. IPCC’s central prediction was accordingly double the observed outturn.

How, then, does the simple model fare? Taking the Planck sensitivity parameter *λ*0 = 0.31 K W–1 m2 and the transience fraction *r*25= 1, and adopting the values given in Table 7 of [6] for the feedback sum *f* and the closed-loop gain *g,* the simple model predicts warming of 0.20 [0.16, 0.26] K since 1990, the central estimate being too low by 0.15 K, the simple model not having taken into account the strong natural forcing during the vigorously positive phase of the Pacific Decadal Oscillation that occurred from 1976 to the turn of the millennium. IPCC’s central estimate, however, was excessive by 0.35 K, well over double the magnitude of the simple model’s error. Table 1 shows the position.

**Table 1**  Comparison of the 1990-2014 transient climate-sensitivity interval 0.20 [0.16, 0.26] K generated by the simple model with the 0.35 K outturn as the linear trend on the mean of the monthly anomalies on three surface and two lower-troposphere datasets and the interval 0.69 [0.48, 1.03] K predicted in FAR on the business-as-usual emissions case.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***f*** | ***g*** | ***λ*∞** | **Δ*F*25** | **Δ*T*25** | **Δ*T*25** | **Δ*T*25** |
|  |  |  |  |  |  |  |
| **[6]** | **[6]** | ***λ*0(1 – *g*)–1** | **0.83 x 5.35 ln**  **(397 / 353)** | **Model**  **(*λ*∞ Δ*F*2x)** | **Mean of**  **5 dsets** | **FAR**  **scen. A** |
| **(W m–2 K–1)** |  | **(K W–1 m2)** | **（W m–2）** | **（K）** | **（K）** | **（K）** |
| –1.60  –0.64  +0.32 | –0.5  –0.2  +0.1 | **0.208**  **0.260**  **0.347** | **0.754** | **0.16**  **0.20**  **0.26** | 0.35 | **0.48**  **0.69**  **1.03** |

The simple model, if the parameter values in [6] are adopted, thus comfortably outperforms the general-circulation models on which IPCC relied in FAR.

Indeed, IPCC – on the advice of expert reviewers including the lead author of the present paper – has realized that the general-circulation models were exaggerating warming, has substituted its “expert assessment” for theirs. It has greatly reduced its medium-term warming projections in AR5 compared with FAR. In 1990, IPCC predicted medium-term warming at 0.3 [0.2, 0.4] K decade–1, but in 2013, following observed warming at 0.14 K decade–1 since 1990, the prediction had been all but halved to just 0.17 [0.10, 0.23] K decade–1. Curiously, though, IPCC has not correspondingly reduced the equilibrium warming projections, which have remained broadly the same since Jule Charney’s influential report of 1979 for the U.S. National Academy of Sciences.

In [13] it is argued that El Niño and the Pacific Decadal Oscillation change the amounts of warming. This is true but irrelevant. It is the purpose and duty of models such as the CMIP5 to incorporate, allow for and represent these and all other atmospheric phenomena related to temperature. Whenever these features are predominant is important, but their presence or absence cannot be prayed in aid to indicate periods during which the CMIP5 models or the model in [6] are valid. In particular, the statistical test that shows that the observed warming during one period differed from that used in [6] is likewise superfluous. The authors of [13] need have said no more than they used values and time points that differed from those in [6].

The plots in Fig. 1 of [13] are not the result of out-of-sample predictions, but comprise model fits to the observed data. The plots are analogous to showing how close a regression line is over a scatter plot of already-observed data. All that these plots prove is that the CMIP5 outputs can be made to fit the data slightly better than those of the simple model. So what? That is what we should expect given that the simple model is “irreducibly simple” and that the CMIP5 ensemble is vastly more complex. And, as the above examples illustrate, the simple model outperforms the general-circulation models over recent decades.

The lower panel in Fig. 1 of [13] shows that the simple model run using different parameter estimates represents the period 1998-present better than with the original parameter estimates (though the fit to the pre-1998 period is then less good). The implication is that the simple model is better than originally thought.

However, all this is a distraction. The real test is always predictive skill. That models should fit observed data well is a necessary but not a sufficient criterion for predictive skill. In [13] the fallacy is perpetrated of assuming that better model fit implies better predictive skill. It does not, as is amply demonstrated in Figs 5-8 here.

Furthermore, RMSE model-fit statistics are generally a poor measure to use in validation because the in-sample model predictions are not univariate numbers, but “envelopes” that should be (but in [13] were not) treated probabilistically. A statistical technique such as the complete-rank probability score is far superior [17].

It is concluded in [13] that the simple model “may not be considered validated”. Here, the authors of [13] make the elementary mistake of assuming that a choice of parameter values that they regard as inappropriate invalidates the model to which that choice of parameters was applied, rather than invalidating solely the chosen parameter values. The simple model, if informed with IPCC’s parameter values, generates climate sensitivities very close to those put forward by IPCC, as the authors of [13] themselves demonstrate, thereby – in their own terms – validating it.

**2.5 Closed-loop gain** In [13] an incomplete summary is given of the grounds in [6] for considering it likely that the closed-loop feedback gain factor *g* < 0.1 (misdescribed in [13] as the “system gain”, which is the term usually applied – and applied in [6] – to the open-loop gain factor *G*). In fact, there is now extensive literature, some of it cited in [6], indicating on the basis of empirical evidence as well as theory that temperature feedbacks are not likely to be strongly net-positive. The parameters in [6] were chosen on the basis of the literature cited, and on the basis of the manifest inapplicability of the Bode system-gain equation to the climate object. The authors of [13] are, of course, as free as anyone else to choose their own parameter values.

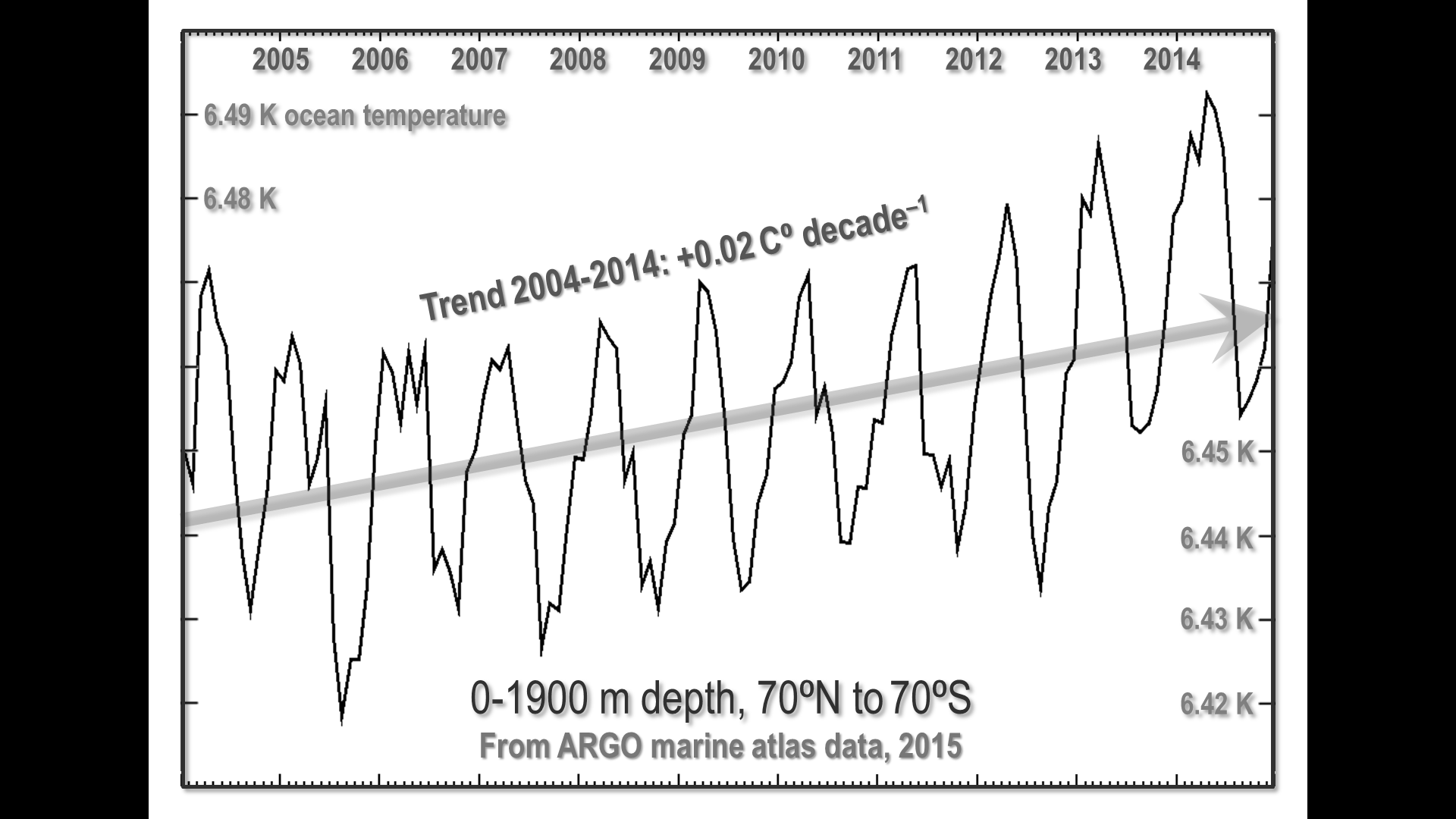
**2.6 Determination of the feedback sum from paleoclimate evidence** The authors of [13] cite [20], whose lead author they misspell as “Zachlos”, as contradicting the assertion in [6], following the cryostratigraphic temperature reconstruction in [21], that paleoclimate temperature varied by only 3.5 K either side of the mean over the past 810 ka. They say [20] found that during the late Paleocene thermal maximum 55 Ma ago temperatures in high latitudes rose by 8 K, perhaps as a result of forcing from greenhouse-gas increases.

However, in [6] it was stated, following [21], that paleoclimate temperature varied by only 3.5 K either side of the mean only over the past 810 ka, not over the past 55 Ma. From the paleocene-eocene thermal maximum to the commencement of the currently-available cryostratigraphic temperature reconstruction, there were substantial and very significant changes in tectonic and continental configurations inconsistent in fundamental respects with today’s climatic environment.

Furthermore, the quotation from [20] given in [13] is incomplete: the words “*and lesser amounts towards the equator”* were omitted. Here, an error is made that is common among those unfamiliar with the consequences of poleward advection, one of the fundamental and often underappreciated non-radiative transfers in the climate object [22-23]. Cryostratigraphy reconstructs temperatures at or near the Poles; but advection from the tropics approximately doubles at the Poles any change in mean global temperatures. This process is known as “polar amplification”. Thus, a 7 K global glacial-to-interglacial interval of temperature change demonstrated in [21] is equivalent to a 14 K change at the Poles. The temperature change of 8 K at the Poles even during the extreme and rapid events of the late Paleocene thermal maximum is thus well within the interval implicit in the cryostratigraphic record presented by [21]. Furthermore, in [20] it is made quite plain that the cause of that thermal maximum, whose onset may have occurred over as little as 1000 years, is unknown. It is not thought that CO2 had anything to do with it, though [20] canvasses the possibility of a substantial dissociation and subsequent oxidation of 2000-2600 Gt of methane from seabed clathrates.

In [13] it is also stated that, if temperature feedbacks were negative, small initial forcings owing to the Milankovic cycles would not be capable of triggering glacial-interglacial transitions. Yet two suggested possible causes of the late Paleocene thermal maximum, discussed in [20], are the large forcings from abrupt deep-sea warming resulting from sudden changes in ocean circulation [24-25] and massive regional slope failure [26].

**2.7 Ocean heat content** In [13] it is asserted that ocean heat content is measured by the ARGO bathythermo-graph buoys in Watts per square meter and that a “net heating” of about 0.5 W m–2 was shown by the buoys from 2000-2010. In fact, [27], the paper cited by the authors of [13], inferred the energy imbalance from satellite observations and used the measurements from the ARGO buoys to determine whether there was an inconsistency between the indications from the two systems. The ARGO buoys do not measure ocean heat content but ocean temperature and salinity, though some published ocean datasets, such as that of NOAA, convert the temperature measurements to ocean heat content. Ocean heat content is in any event not measured in Watts per square meter but, typically, in multiples of 1022 Joules. And Watts per square meter are a derivative SI unit measuring incident radiation at a surface and do not measure “heating”, which is measured in Kelvin, a base SI unit. Furthermore, the buoys did not have sufficient spatial coverage to begin to supply near-global data till 2004, so that the first four years of the 11-year period mentioned in [13] were not global measurements at all.

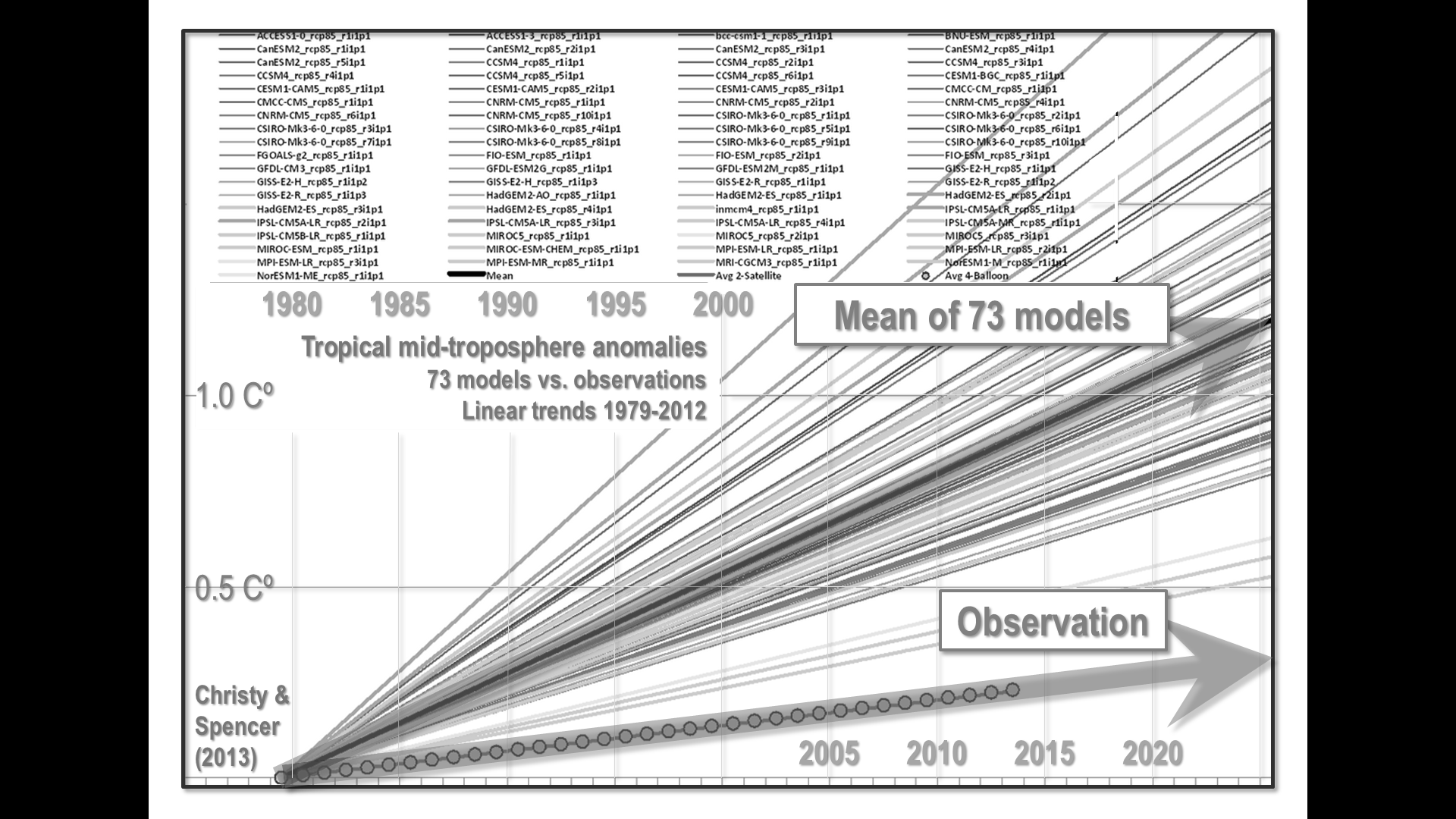


**Fig. 4** Ocean temperature, 0-1900 m depth, 70ºN to 70ºS, over the entire ARGO time-series from 2004-2014. Since the ARGO buoys measure temperature, the ARGO data from 2004-2014 were obtained from the ARGO marine atlas and plotted in Fig. 4, together with the linear trend. The data (albeit subject to very substantial uncertainties in that each buoy must attempt to monitor and represent the temperature changes in 200,000 km2 of ocean) indeed demonstrate warming of the ocean over the 11-year period for which data are available. The warming trend over the decade was 0.02 K. If this trend were the consequence of 0.5 W m–2 of radiative forcing, then the implicit transient-sensitivity parameter would be 0.04 K W m–2. Even after allowing for thermal inertia in the oceans, the implication is that strongly net-negative temperature feedbacks are in operation, confirming, contrary to an assertion in [13], that little error arises from assuming *rt* = 1

**2.8 Values of the transience fraction *rt***  In [13] it is argued that the choice of *rt* = 1 in [6] was “equivalent to an instantaneous response”, but that “this is only true if the heat capacity of Earth is zero”. Here, the authors of [13] have misunderstood the definition of the transience fraction. In [6], the transience fraction was defined as “the fraction of equilibrium sensitivity expected to be attained over *t* years”, but, in the discussion of this variable, it was made plain that its purpose was to allow for non-linearities specifically arising from the action of temperature feedbacks over different timescales. There are indeed, therefore, response lags in the climate object caused by the immense thermal inertia of the ocean, for instance: but the simple model concerns itself only with anthropogenic influences on temperature. Thermal inertia is not an anthropogenic influence.

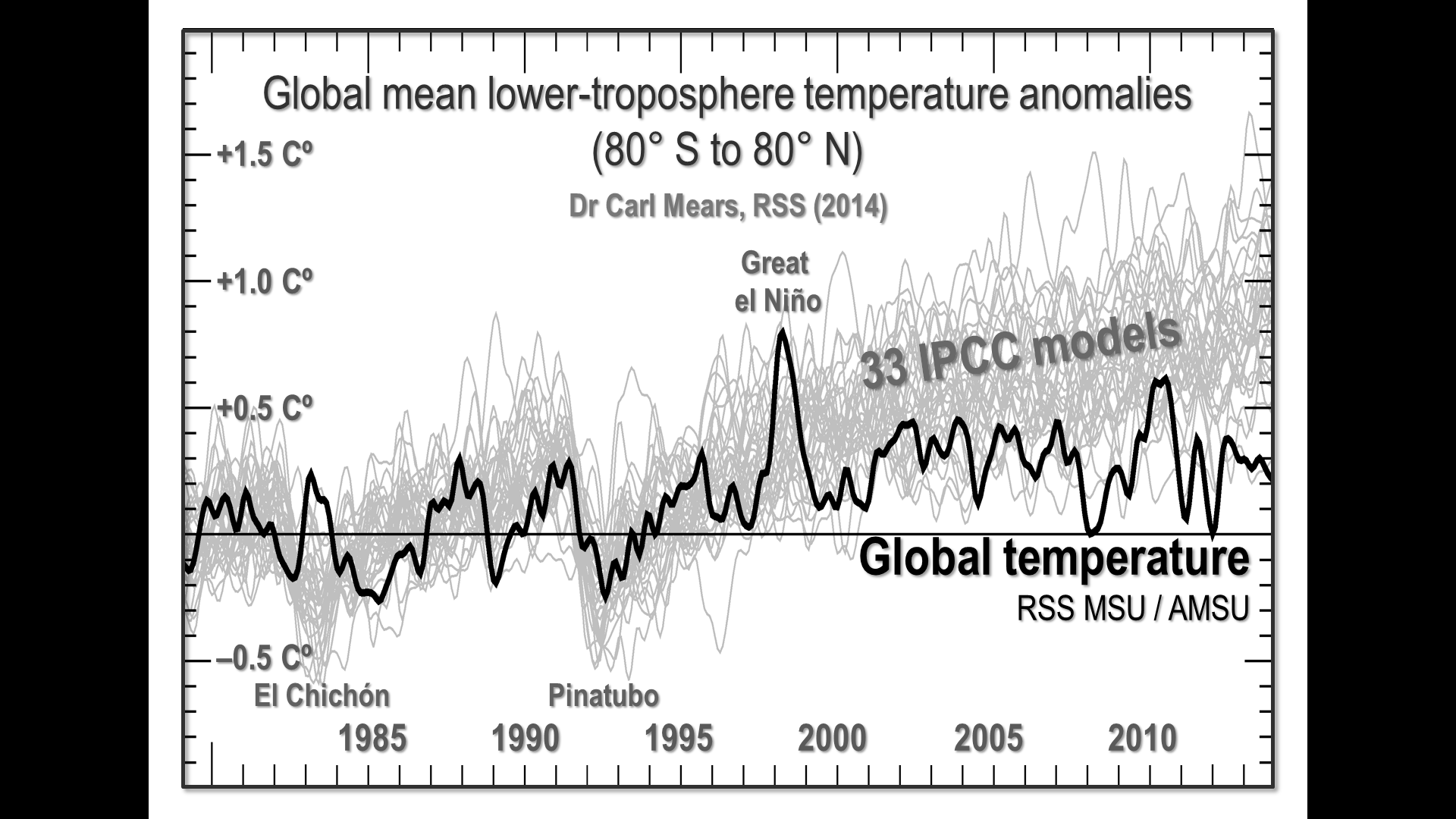
**2.9 Reasons why the general-circulation models’ predictions have proven excessive** In [13] it is stated that “temperature trends since approximately 1998 have been at the low end of the CMIP5 distribution”. In fact, it is not appropriate to consider only the period since 1998, for nearly all of that period falls within the current negative phase of the Pacific Decadal Oscillation, whose 30-year cycles of positive or warming phases followed by 30-year cycles of cooling phases have had a visible effect on the evolution of global temperature since the beginning of the 20th century. These cycles must be allowed for in attribution of temperature trends. The simplest method is to ensure that any period of study is either a multiple of 60 years or, if less than 60 years, is centered on a phase-transition between the PDO phases. Conveniently, IPCC made its first temperature predictions in 1990 and almost half of the period since then was in a positive PDO phase, the remainder being in a negative phase. As fig. 1 of [6] demonstrated, IPCC’s central medium-term global-warming prediction in 1990 has proven to be double the observed warming rate.

The failure of the models has been sufficiently severe to persuade IPCC itself, in AR5, to substitute its own “expert assessment” for their medium-term output. A measure of the severity of the models’ failure is shown in Fig. 5, where the linear trends on the influential tropical mid-troposphere temperatures as predicted by 73 models are compared with the linear trends as observed and recorded in two global-temperature datasets.



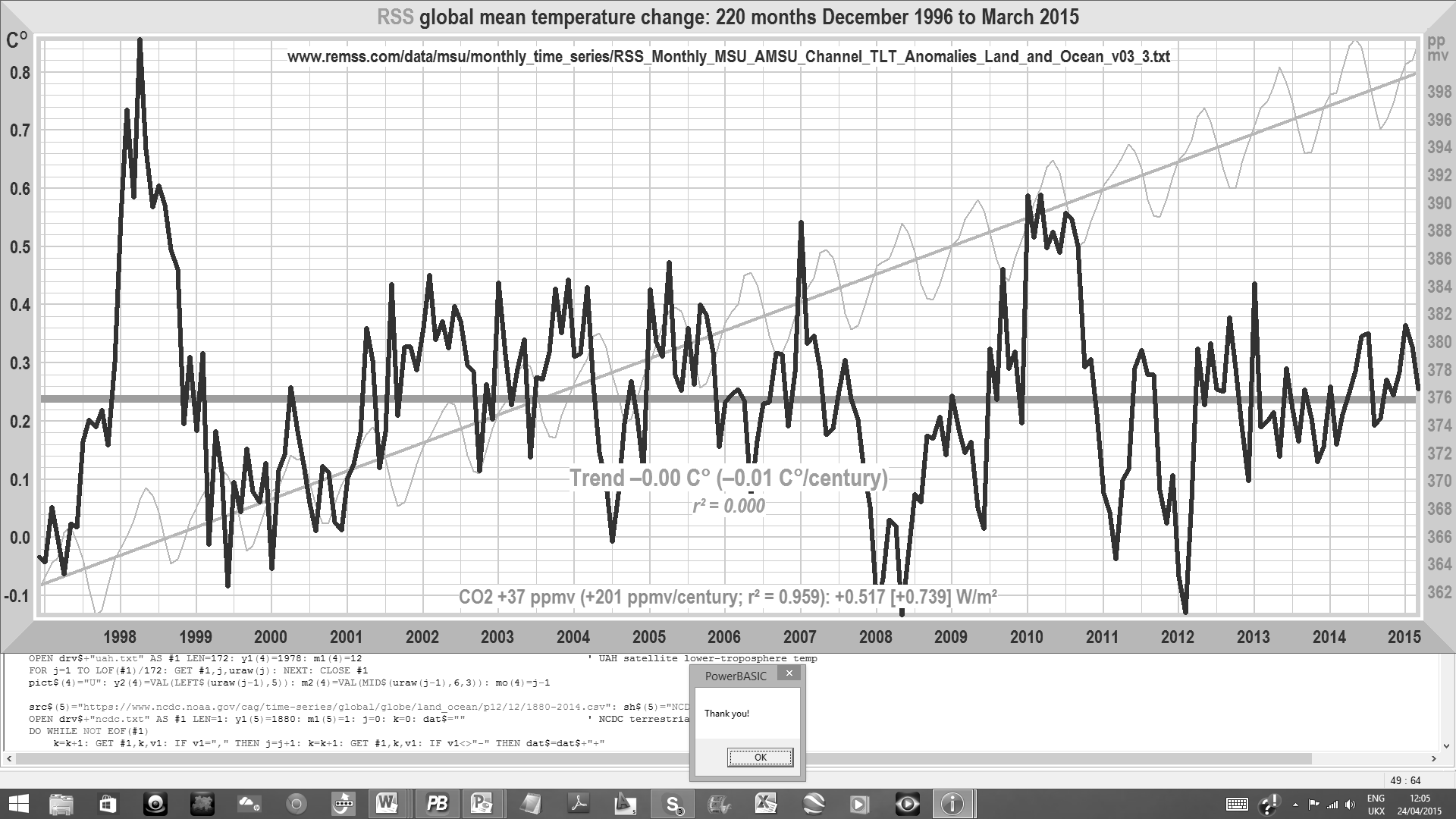
**Fig. 5** Linear trends on tropical mid-troposphere temperature anomalies projected by 73 models and measured by two coincident observational datasets (UAH and balloon radiosondes), 1979-2012. The mean linear trend on the 73 models is the upper shaded arrow

Likewise, the RSS satellite dataset indicates a considerable exaggeration of the warming rate by the models compared with observation (Fig. 6).



**Fig. 6** Mean lower-troposphere temperature anomalies, 80º S to 80º N, 1979-2014 (black spline-curve) against the anomalies predicted by 33 models (based on Mears, 2014)

It is recognized in [13] that the general-circulation models have failed in the task of predicting global temperature change. Indeed, three possible explanations are offered: overestimated radiative forcing; inadequate representation of internal variability; and overestimated climate response. It is suggested that the authors of [6] favor the third explanation of the models’ running hot. However, the state of knowledge of the magnitude of radiative forcings and of the mechanisms of internal climate variability is insufficient to allow any definite conclusion to be drawn as to the reasons for the general-circulation models’ failure. As an instance, the startling abruptness and magnitude of the Great El Niño of 1997-1998 (Fig. 7) are not yet satisfactorily explained. Accordingly, [6] concentrated on five specific aspects of the models’ representation of climate that were likely to have contributed significantly to the considerable overstatements of warming in the models’ projections.



**Fig. 7** RSS monthly global mean lower-troposphere temperature anomalies over the 220 months December 1996 to March 2015. Over these 18 years 4 months, the linear trend on the anomalies is approximately zero, notwithstanding CO2 concentration (pale gray curve and trend) rising at a rate equivalent to >200 ppmv century–1. The aetiology of the Great El Niño of 1997-1998, which according to the evidence of coral bleaching had two precedents in the previous 300 years, is not yet satisfactorily explained

The second satellite dataset, UAH, shows a similar period of temperature stasis. None of the complex models predicted anything like so long a hiatus in global warming. The discrepancy between prediction and observation over that 18-year period is smaller for the simple model than for the general-circulation models.

**2.10 IPCC’s estimates of temperature feedbacks** It is argued in [13] that [6] had misinterpreted IPCC feedback estimates, in that the magnitudes of the principal climate-related feedbacks had been determined by different methods in AR4 and AR5, the latter having included representations of ocean changes and also having assumed a linear feedback response, and in that a second panel in Fig. 9.43 of AR5 had not been reproduced alongside the first. Here, the authors of [13] are more than somewhat disingenuous. Fig. 9.43 of AR5 was visibly intended to provide a direct comparison between the magnitudes of temperature feedbacks in AR4 and AR5. The feedback sum in AR5 is substantially less than that in AR4, because the more complete representation of the coupled ocean-atmosphere system in AR5 has constrained the feedback values (which, however, remain substantially overstated). Furthermore, the second panel of Fig. 9.43 was omitted in [6] because it did not concern itself with individual feedback values; nor did it compare the feedback sums in AR4 and AR5.

The stated purpose of Fig. 9.43 in AR5 was to provide a direct comparison between the feedback values used by the CMIP3 models in AR4 and the CM5 models used in AR5. For the reasons explained in [6], the feedback values in AR4 lead to the climate sensitivities found in AR4, but the same is not the case for AR5, where climate sensitivity should have been – but was not – reduced by one-third to take account of the reduction in feedback values. Contrary to an assertion in [13], neither the text of AR5 nor the missing panel from Fig. 9.43 in any way explains the reasons why a substantial reduction in feedback values in response to an unchanged forcing can lead to an unchanged climate sensitivity.

As for the fact that AR5 assumed a linear feedback response, the implications of this assumption for climate sensitivity are by no means clear, not least because the values of individual feedbacks, as well as the curves of their evolution over time, are so uncertain. The Appendix to [6], which was lost for reasons of space, is annexed here to outline the mathematical treatment of non-linear feedbacks.

**3 Discussion**

In [6], the simple model provided a test bench against which some of the assumptions and methods in the general-circulation models could be tested, and five defects in those models were identified and discussed. In [13], three of the five conclusions of the experiment in [6] were challenged, but, for the reasons described above, the challenges were not compelling.

Two of the five conclusions in [6] were not challenged at all in [13]: that there is no global warming “in the pipeline” in consequence of our past emissions, and that the extreme RCP 8.5 scenario is unjustifiable. We are gratified that the authors of [13] seem to have found common ground with us on these points.

In [13], an attempt was made to demonstrate that the predictive skill of the simple model presented in [6] was worse than that of the general-circulation models. However, that attempt merely served to confirm the skill of the simple model in replicating IPCC’s climate sensitivities if IPCC’s values for the model’s parameters were adopted. And tests of the model against observed temperature change show that since 1750 the simple model performs much as the general-circulation models perform, and since 1990 the model has performed with very much greater skill than the general-circulation models.

Very nearly all of the attempted criticisms of [6], though presented as criticisms of the simple model, were in fact directed solely at the choice of parameter values, and not of the simple model itself, which the authors of [13] inadvertently validated in accordance with their own methodology by adopting parameter values consistent with IPCC reports and consequently obtaining IPCC estimates of climate sensitivity from the simple model.

One point that emerges from comparison of the relative merits and skills of the simple and general-circulation models is the sheer magnitude of the influence of uncertainty on any attempt to project future climate states. IPCC, to its credit, has consistently conceded that uncertainties in cloud and aerosol forcings and feedbacks are substantial. The recent re-evaluation of the impact of anthropogenic aerosols on the determination of climate sensitivity in [17] is a case in point. FAR had made little allowance for aerosols. Subsequent IPCC *Assessment Reports* made so much allowance for them that forcings since the industrial revolution were all but halved. This decision had the effect of greatly increasing estimates of climate sensitivity derived from observation of temperature change since 1750. If [17] is correct, however (the literature is strongly divided on the aerosol question), then perhaps FAR was right all along about the magnitude of post-1750 forcings.

IPCC has also recognized, correctly, that the climate behaves as a coupled, non-linear, chaotic object and that, accordingly, the long-term prediction of future climate states is not possible [TAR, §14.2.2.2; cf. 28-29].

The greatest source of uncertainty in determining climate sensitivity lies in the temperature feedbacks. That feedbacks exist is self-evident. However, their magnitudes and even in some instances their signs are unknown. Here, IPCC and the general-circulation models are at fault in assuming that the magnitudes of most temperature feedbacks are well constrained. Plainly, this is not the case, as the reduction in the feedback-sum *f* from 2.0 in AR4 to 1.5 in AR5 demonstrates. For the reasons discussed in [6], it is likely that the feedback-sum is net-negative, whereupon climate sensitivity cannot exceed an upper bound of 1.3 K per CO2 doubling.

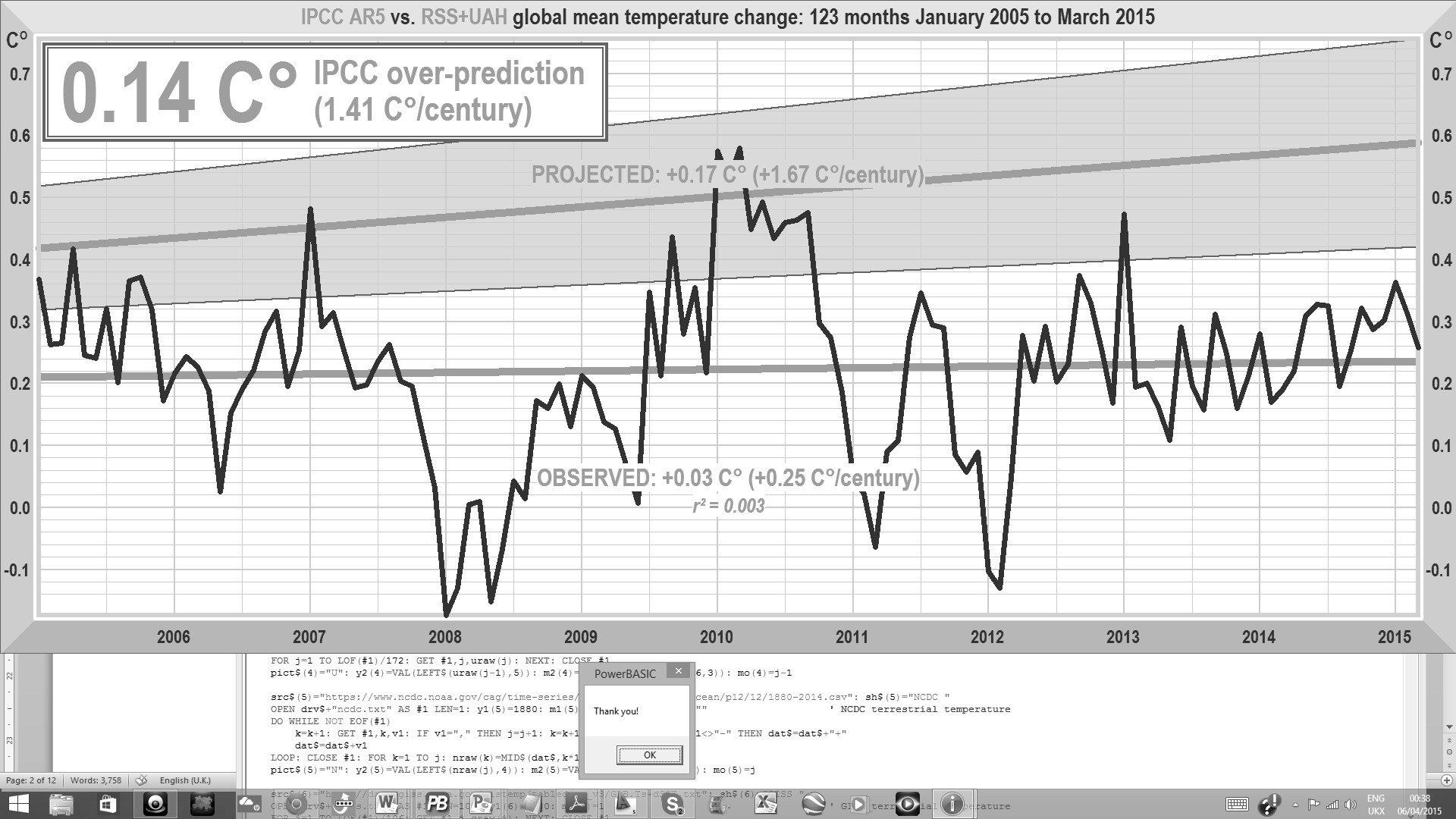
The atmosphere, sandwiched between two vast heat-sinks – the ocean below and outer space above – has proven formidably but unsurprisingly thermostatic over the past 810,000 years [21]. In that circumstance, the only way to maintain that climate sensitivity is high is to maintain – incorrectly, as evidenced in [20, 22-23] – that no large forcings occurred in the later Pleistocene and Holocene epochs and that, therefore, the small temperature change of the past 810,000 years was attributable to large feedbacks operating on small forcings.

Finally, we wish to note that, though it is falsely claimed in [13] that our cited references [7-8] have been “collectively rebutted”, no specific arguments or references were provided. Such arguments, in fact, have *not* been rebutted but rather have become accepted by mainstream science. For example, in the case of the infrared iris proposed by the lead author of [7], at least two recent publications [30-31] suggest findings consistent with the hypothesis.

**4 Conclusions**

The general-circulation models now face a crisis of credibility. Not one of them predicted a stasis of as long as 17-18 years in global temperatures. Indeed, it is often stated that periods ≥15 years without warming are inconsistent with models’ predictions. For instance, [32-33] state: “The simulations rule out (at the 95% level) zero trends for intervals of 15 yr or more, suggesting that an observed absence of warming of this duration is needed to create a discrepancy with the expected present-day warming rate.” See also [34-36].

The models relied upon in FAR predicted twice as much warming from 1990-2014 as has been observed. All models predicted a warming rate in the crucial tropical mid-troposphere considerably in excess of observation. It is no longer credible to ignore these ever-widening discrepancies between prediction and observation. IPCC itself has recognized that, at least as far as medium-term prediction is concerned, the models have failed, raising the legitimate question whether the longer-term predictions have also been exaggerated, perhaps greatly. The best estimate in FAR was that in the 40 years 1990-2030 global temperature would rise at an approximately linear rate by 1.8 K. After 25 of those 40 years, just 0.35 K global warming has occurred – a rate equivalent to 0.14 K decade–1. To reach 1.8 K in the next 15 years, the warming rate would have to accelerate to almost 1 K decade–1, or seven times the rate observed over the past 25 years. Recently [36] drew attention to the fact that internal variability may play an important role in explaining the apparent near-zero trend in September Arctic sea-ice extent record from 2007-2013.



**Fig. 8** Medium-term temperature projections in AR5 (gray region) against the mean of the RSS & UAH monthly global mean lower-troposphere temperature anomalies and the trend thereon

IPCC, unlike the authors of [13], has taken note of the failure of the general circulation models’ predictions and has not sought to maintain they were accurate when they were not. In AR5 IPCC cut its central prediction of medium-term warming from 0.29 to 0.17 K decade–1, or little more than the observed rate of 0.14 K decade–1 over the past 25 years. However, even that revised prediction is proving excessive (Fig. 8).

Over the past decade, the over-prediction even on the new, lower rate of predicted medium-term warming has proven to be 0.14 K. For the measured atmospheric warming over the period was equivalent to 0.25 K century–1, consistent with the underlying ocean warming rate of 0.23 K century–1 in the ARGO monthly anomalies over a very similar period (Fig. 4).

In the circumstances, it cannot be safely said that, at least as to transient climate sensitivity, the general-circulation models are better validated than the simple model. See [37-39] for a fuller discussion of model validation, which is a more complex task than simply the analysis of RMSE and bias mentioned in [13].

The value of the simple model, particularly when informed by parameters arguably more reasonable than those adopted by the IPCC, is in facilitating examination of the reasons why such extreme over-predictions are being made – predictions that are at present being acted upon by governments.

As to whether the predictions of general-circulation models or of the simple model will prove correct in the long run, only time will tell. For recent decades, though, the simple model is proving closer to reality.

The pressing question arises why the general-circulation models’ central longer-term projections suggest as much as 3 K warming by 2100 compared with today, let alone why still more extreme estimates of up to 6 K 21st-century warming are made in some circles. The simple model indicates that high sensitivity is implausible.

**Appendix 1: Further development of the model**

The model may readily be further developed to increase its sophistication, though such developments are beyond the scope of the present paper. For instance, an additional factor might be included in (1) to represent any desired contribution from anthropogenic forcings.

The model might also be made one-dimensional, by representing the latitude. One-dimensional energy-balance models (see [40] for an overview), originally developed by [41-42] and extended by [43-44], have been widely used to introduce students to climate modeling and to examine some peculiarities of the climate system. Some of the more interesting issues that appear when latitude is taken into account are polar amplification of sensitivity to a forcing, the snowball/snow-free bi-stability [40] and the small ice-cap instabilities [45] that arise from the positive ice-albedo feedback.

A one-dimensional model starts with (1) and, at each latitude *φ*, expresses the albedo *α* of the Earth and its clouds, its effective temperature *TE*, and the distribution of solar irradiance *S* as functions of *x =* sin *φ*. The one-dimensional model also implicitly assumes that the northern and southern hemispheres are reflections of one another with no net heat flux across the equator. To resolve the latitudinal dimension, the model may be grid-based (as in [43-44]) or based on Legendre polynomials (as in [46]).

The model, however formulated, requires a further equation to describe the meridional or poleward transfer of heat. This energy-flux divergence *D* is proportional to –2T. Using Fick’s Law of Diffusion in (A1.1), it is expressed by

, | *x =* sin *φ* , (A1.1)

where the diffusion coefficient, , representing the poleward transfer of energy via oceanic and atmospheric advection, is a tunable parameter that yields a realistic equator-to-pole temperature gradient and can be simplified to render it independent of latitude, with a customary value ~0.65 W m–2 K–1. In [47] a diffusion coefficient is suggested that is dependent on *x2* (consistent with the diffusion coefficient in [41]) so that tropically-averaged motions are better described,

, (A1.2)

where is the second Legendre polynomial and , are tunable parameters [see also 48]. Use of the second Legendre polynomial is fortunate in that should decrease toward the pole, as it does in (A1.2), and the poleward heat flux vanishes at the Equator.

Given the complicated motions of the atmosphere and ocean that transport the energy poleward, such diffusive approximations are conceptually appealing but may not be entirely physically-based [46]. Formulation of diffusion using (8) introduces a term that varies as a function of the equator-to-pole temperature distribution which, necessarily, will alter the temperature response Δ*Tt* to anthropogenic radiative forcings, thereby changing the response in (1). Specifically, addition of latitudinal diffusion will affect not only the transience fraction, *rt*, since the impact of diffusive heat transport and its response to Δ*T* will change the response time to anthropogenic forcing, but also the equilibrium climate-sensitivity parameter, *λ*∞.

The model may also be developed to represent non-linear temperature feedbacks. Where feedbacks are non-linear (see [49] for the derivation), (4) becomes (A1.3):

. (A1.3)

In the general case, therefore, the linear-feedback system-gain relation *Gt =* (1 – *gt*) –1 becomes (A1.4):

| *ξ* = 0 where feedbacks are linear. (A1.4)

**Ethical declaration**

The authors declare that they have no conflict of interest.

**References**

1. IPCC FAR (1990) Climate change – The IPCC Assessment (1990): **Report prepared for Intergovernmental Panel on Climate Change by Working Group I. In: Houghton JT,** Jenkins GJ, Ephraums JJ (eds) Cambridge University Press, Cambridge, New York, Melbourne
2. IPCC SAR (1995) Climate Change 1995 – The science of climate change: contribution of WG1 to the second assessment report. Houghton JT, Meira Filho LG, Callander BA et al (eds) Cambridge University Press, Cambridge, New York, Melbourne
3. IPCC TAR (2001) Climate Change 2001: The scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change. Houghton JT, Ding Y, Griggs DJ, et al (eds) Cambridge University Press, Cambridge, New York, Melbourne
4. IPCC AR4 (2007) Climate change 2007: The physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change, 2007. In: Solomon S, Qin D, Manning M, et al (eds) Cambridge University Press, Cambridge, New York, Melbourne
5. IPCC AR5 (2013) Climate change 2013: The physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. In: Stocker TF, Qin D, Plattner G-K, et al (eds) Cambridge University Press, Cambridge, New York, Melbourne
6. Monckton of Brenchley, CW, Soon WW-H, Legates DR et al (2015) Why models run hot: results from an irreducibly simple model. Sci Bull 60:122-135, doi:10.1007/s11434-014-0699-2.
7. Lindzen, RS, Choi Y-S (2011) On the observational determination of climate sensitivity and its implications. Asia-Pacific J Atmos Sci 47:377–390
8. Spencer RW, Braswell WD (2011) On the misdiagnosis of surface temperature feedbacks from variations in Earth’s radiant-energy balance. Remote Sensing 3:1603–1613
9. Monckton of Brenchley, CW (2015) The temperature feedback problem. Energy & Environ.[in press]
10. RSS (2015) Satellite-derived monthly global mean lower-troposphere temperature anomaly dataset: www.remss.com/data/msu/monthly\_time\_series/RSS\_Monthly\_MSU\_AMSU\_Channel\_TLT\_Anomalies\_Land\_and\_Ocean\_v03\_3.txt. Accessed 4 April 2015
11. NOAA (2015) Atmospheric CO2 concentration expressed as a mole fraction in dry air (μmol mol–1) ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2\_mm\_mlo.txt. Accessed 29 April 2015
12. Myhre G, Highwood EJ, Shine KP et al (1998) New estimates of radiative forcing due to well-mixed greenhouse gases. Geophys Res Lett 25:2715–2718
13. Richardson M, Hausfather Z, Nuccitelli D et al (2015) Misdiagnosis of Earth climate sensitivity based on energy balance model results. Sci Bull [in press].
14. Connolly R, Connolly M (2014a) Urbanization bias I: is it a negligible problem for global temperature estimates? Open Peer rev J 28 (Clim Sci) ver 0.1, oprj.net/articles/climate-science/28 Accessed 27 December 2014
15. Connolly R, Connolly M (2014b) Urbanization bias III: estimating the extent of bias in the Historical Climatology Network datasets. Open Peer rev J 28 (Clim Sci) ver 0.1, oprj.net/articles/climate-science/34 Accessed 27 December 2014
16. Le Quéré C, Moriarty R, Andrew RM et al (2014) Global carbon budget 2014. Earth Syst Sci Data Discuss 7:521-610, doi:10.5194/essdd-7-521-2014
17. Gneiting T, Raftery AE (2007) Strictly Proper Scoring Rules, Prediction, and Estimation. JASA 102:359-378
18. Lewis N (2015) The implications for climate sensitivity of Bjorn Stevens’ new aerosol forcing paper, http://climateaudit.org/2015/03/19/the-implications-for-climate-sensitivity-of-bjorn-stevens-new-aerosol-forcing-paper, accessed May 2, 2015
19. Stevens B (2015) Rethinking the lower bound on aerosol radiative forcing. J. Clim., doi:10.1175/CLI-D-14-00656.1
20. Zachos J, Pagani M, Sloan L et al (2001) Trends, rhythms and aberrations in global climate 65 Ma to present. Science 292:686–693
21. Jouzel J, Masson-Delmotte V, Cattani O et al (2007) Orbital and millennial Antarctic climate variability over the past 800,000 years. Science 317:793–796
22. Lindzen RS (1994) Climate dynamics and global change. *Ann Rve Fluid Mech,* 26:353-378
23. Soon WW-H, Legates DR (2013) Solar irradiance modulation of equator-to-pole (Arctic) temperature gradients: empirical evidence for climate variation on multi-decadal timescales. J Atm Solar-Terr Phys 93:45-56, doi:10.1016/j.jastp.2012.11.015
24. Kaiho K, Arinobu T, Ishiwatari R et al (1996) Latest Paleocene benthic foraminiferal extinction and environmental changes at Tawanui, New Zdaland. Paleooceanography 11:447, doi: 10.1029/96PA01021
25. Bralower TJ, Thomas DC, Zachos JC et al (1997) High-resolution records of the late Paleocene thermal maximum and circum-Caribbean volcanism: Is there a causal link? Geology 25(11):963-966
26. Katz ME, Pak DK, Dickens GR et al (1999) The source and fate of massive carbon input during the latest Paleocene thermal maximum, Science 286(5444):1531-1533
27. Loeb, NG, Lyman JM, Johnson GC et al (2012) Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. Nature Geosci 5:110-113, doi: 10.1038/ngeo1375
28. Lorenz EN (1963)Deterministic nonperiodic flow.J Atmos Sci 20:130-141.
29. Giorgi F (2005)Climate Change Prediction.Clim Change 73:239-265, doi:10.1007/s10584-005-6857-4
30. Mauritsen T, Stevens B (2015) Missing IRIS effect as a possible cause of muted hydrological change and high climate sensitivity in models. Nature Geosci. [in press]
31. Tan J, Jakob C, Rossow WB et al (2015) Increases in tropical rainfall driven by changes in frequency of organized deep convection. Nature 519:451–454
32. Peterson TC, Baringer MO [Eds.] (2009) State of the climate in 2008. *Bull Am Met Soc* 90:8, Supplement
33. Knight J, Kennedy JJ, Folland C et al (2009) Do global temperature trends over the last decade falsify climate predictions? Bull Am Met Soc 90:S22-S23
34. Santer BD, Mears C, Doutriaux C et al (2011) Separating signal and noise in atmospheric temperature changes: The importance of timescale. J Geophys Res, doi:116:2011JD016263
35. Fyfe JC, Gillett NP, Zwiers FW (2013) Overestimated global warming over the past 20 years. Nature Clim Change 3:767-769
36. Swart NC, Fyfe JC, Hawkins E et al (2015) Influence of internal variability on Arctic sea-ice trends. Nature Clim Change 5:86-89
37. Legates DR, McCabe CJ Jr (1999) Evaluating the use of “goodness of fit” measures in hydrologic and hydroclimatic model validation. Water Resources Research 35(1):233-241
38. Legates DR, McCabe CJ Jr (2013) A refined index of model performance: a rejoinder”. Int J Climatol, 33(4):1053-1056
39. Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'Donnell J, Rowe CM (1985) Statistics for the evaluation and comparison of models. *J Geophys Res* **90**(C5):8995-9005
40. Bódai T, Lucarini V, Lunkeit F et al (2014) Global instability in the Ghil-Sellers model. Clim Dyn [in press]. doi:10.1007/s00382-014-2206-5
41. Budyko MI (1969) The effect of solar radiation variations on the climate of the Earth. Tellus 21:611–619
42. Sellers WD (1969) A global climatic model based on the energy balance of the earth-atmosphere system. J Appl.Meteorol 8:392–400
43. North GR (1975) Theory of energy-balance climate models. J Atmos Sc. 32:2033–2043
44. Ghil M (1976) Climate stability for a Sellers-type model. J Atmos Sci 33:3–20
45. North GR (1984) The small ice cap instability in diffusive climate models. J Atmos Sci 41:3390–3395
46. North GR, Cahalan RF, Coakley JA Jr (1981) Energy balance climate models. Rev Geophys Space Phys 19:91–121
47. Lindzen RS, Farrell B (1977) Some realistic modifications of simple climate models. J Atmos Sci 34:1487–1501
48. Schneider EK, Lindzen RS (1977) Axially symmetric steady-state models of the basic state for instability and climate studies, Part I, Linearized calculations. J Atmos Sci 34:263–279
49. Roe G (2009) Feedbacks, timescales, and seeing red. Ann Rev Earth Planet Sci 37:93–115