AGU PUBLICATIONS

Earth and Space Science

RESEARCH ARTICLE
10.1002/2015EA000154

Key Points:
- Support for low observational climate sensitivity estimates that use two-zone energy balance models
- Meridional heat transport and varying radiative response can strongly influence sensitivity estimates
- Sensitivity-altering climate feedbacks are not always additive

Supporting Information:
- Supporting Information S1

Correspondence to:
J. R. Bates,
ray.bates@ucd.ie

Citation:
Bates, J. R. (2016), Estimating climate sensitivity using two-zone energy balance models, *Earth and Space Science*, *3*, doi:10.1002/2015EA000154.

Received 23 DEC 2015
Accepted 26 APR 2016
Accepted article online 6 MAY 2016

Estimating climate sensitivity using two-zone energy balance models

J. Ray Bates[1]

[1]Meteorology and Climate Centre, School of Mathematics and Statistics, University College Dublin, Dublin, Ireland

Abstract Estimates of $2 \times CO_2$ equilibrium climate sensitivity (EqCS) derive from running global climate models (GCMs) to equilibrium. Estimates of effective climate sensitivity (EfCS) are the corresponding quantities obtained using transient GCM output or observations. The EfCS approach uses an accompanying energy balance model (EBM), the zero-dimensional model (ZDM) being standard. GCM values of EqCS and EfCS vary widely [Intergovernmental Panel on Climate Change range: (1.5, 4.5)°C] and have failed to converge over the past 35 years. Recently, attempts have been made to refine the EfCS approach by using two-zone (tropical/extratropical) EBMs. When applied using satellite radiation data, these give low and tightly constrained EfCS values, in the neighborhood of 1°C. These low observational EfCS/two-zone EBM values have been questioned because (a) they disagree with higher observational EfCS/ZDM values and (b) the EfCS/two-zone EBM values given by GCMs are poorly correlated with the standard GCM sensitivity estimates. The validity of the low observational EfCS/two-zone EBM values is here explored, with focus on the limitations of the observational EfCS/ZDM approach, the disagreement between the GCM and observational radiative responses to surface temperature perturbations in the tropics, and on the modified EfCS values provided by an extended two-zone EBM that includes an explicit parameterization of dynamical heat transport. The results support the low observational EfCS/two-zone EBM values, indicating that objections (a) and (b) to these values both need to be reconsidered. It is shown that in the EBM with explicit dynamical heat transport the traditional formulism of climate feedbacks can break down because of lack of additivity.

1. Introduction

A major current issue in climate science is that, despite the considerable progress that has been made in the development of global climate models (GCMs), it has not proven possible to narrow the uncertainty limits of $2 \times CO_2$ equilibrium climate sensitivity since the *Charney Report* [1979]. Recently, it has been seen necessary to widen these limits, with *Intergovernmental Panel on Climate Change* (*IPCC*) [2013] giving a probable range of (1.5, 4.5)°C as compared with the *IPCC* [2007] range of (2, 4.5)°C.

In *IPCC* [2013] the *equilibrium climate sensitivity* refers to the equilibrium (steady state) change in the annual global-mean surface temperature following a doubling of the atmospheric equivalent CO_2 concentration. It is assumed that the real climate system has a well-defined value of equilibrium climate sensitivity; this will here be denoted ECS.

Estimates of the ECS can be obtained by running GCMs to equilibrium. In this case, the GCMs used must be atmosphere-ocean GCMs (AOGCMs). The most comprehensive category of AOGCM involves a dynamic ocean. Because of computational constraints, this is sometimes replaced by a more restricted category of AOGCM involving a mixed layer ocean. In this paper, estimates of the ECS obtained by integrating either of the above categories of AOGCM to equilibrium will be denoted EqCS.

Estimates of the ECS can also be obtained using observations or transient GCM output, in conjunction with a simple energy balance model (EBM) of the climate system. In this case, if a GCM is used, it can be either an AOGCM or a GCM run with observed sea surface temperature (SST) and sea ice extent (the latter referred to as an Atmospheric Model Intercomparison Project (AMIP) GCM; see Appendix A for definitions). *IPCC* [2013] uses the term *effective climate sensitivity* to refer to estimates of the ECS obtained in this way. In this paper, estimates of the ECS so obtained will be denoted EfCS. We shall not be concerned in this paper with the quantity denoted *transient climate response* in *IPCC* [2013; defined in Annex III, p. 1451].

The dominant category of EBM used in the EfCS context is the simple zero-dimensional model (ZDM), in which all quantities are global and annual means and the surface temperature (T) is the basic climatic variable. Perturbations in T around its observed equilibrium value lead to perturbations in the top of the atmosphere (TOA) radiation budget. In the context of EBMs, the radiative perturbation is normally expressed in linearized form as a product of the perturbation in T and a radiative response coefficient, commonly called a feedback parameter (or thermal damping rate) in the climate literature.

In the ZDM, the horizontal heat transport convergence integrates to zero, so that any possibility of perturbations in dynamical heat transport (DHT) influencing EfCS is hidden. Similarly, the ZDM has only a single global mean radiative response coefficient, so that there is no explicit reference to the possibility of interregional differences in this coefficient influencing EfCS.

The current dominance of the ZDM in conceptual discussions of climate sensitivity and the processes that influence it can be seen by reference to *IPCC* [2013], sections TS.5.3, 1.2.2, 7.2.4, 7.2.5, 9.7, 10.8, and Box 12.2. There, the conceptual discussion is based almost entirely on the ZDM framework. The processes currently recognized as influencing climate sensitivity are summarized in section 1.2.2 and illustrated in Figure 1.2 of the IPCC report; perturbations in the atmospheric component of DHT do not appear in Figure 1.2, although the figure does contain a reference to "ocean circulation." In the annual mean poleward heat transport out of the band 30°S–30°N, atmospheric transport is the dominant contributor, comprising about 80% of the total [*Trenberth and Caron*, 2001, Figure 7]. This suggests that it also dominates DHT perturbations.

In observational and GCM studies of EfCS, the ZDM is similarly dominant [see for example, *IPCC*, 2013, Box 12.2, Figure 1]. In the GCM context, the EfCS/ZDM method introduced by *Gregory et al*. [2004] has in recent years gained a central place [see *IPCC*, 2013, section 9.7.1].

In our discussion of EfCS in the present paper, we consider the use of two separate two-zone (tropical/extratropical) EBMs as well as the ZDM. In the first of these, termed Model A, the radiative response coefficients (defined in equation (1)) for the tropics and extratropics are estimated separately, but it is shown that only their average value influences the EfCS; also, there is no option of varying the strength of the perturbation DHT between the zones. Unlike the ZDM, however, Model A implicitly incorporates a perturbation DHT and its value can be determined diagnostically. This model was introduced by Lindzen et al. [*Lindzen et al.*, 2001, hereafter LCH01] and used by *Lindzen and Choi* [2009, 2011; hereafter LC09 and LC11]. These authors recognized that the tropical radiative response coefficient could be estimated to within narrow limits from available observational data, while its extratropical counterpart was difficult to estimate observationally because of the restricted satellite coverage at high latitudes and the predominance of noise in surface temperatures over land. Subsequently, *Choi et al*. [2014] found that fluctuations in SW radiation arising from random internal variations in cloud cover not caused by variations in T dominated the attempted measurements of the SW component of the radiative response coefficient in both zones, leaving this component essentially undetectable. They found that only the LW component in the tropics could be reliably estimated. Their findings are taken into account in the present paper. LCH01, LC09, and LC11 assumed that the radiative response coefficient in the extratropics could be approximated adequately by assuming a blackbody or Planck response. Here we relax that assumption by allowing a range of values of the extratropical quantity.

In our second two-zone EBM, termed Model B, the tropical and extratropical radiative response coefficients are again estimated separately, but perturbations in DHT between the zones are explicitly parameterized and of adjustable strength. The stability and sensitivity properties of this model have been studied by *Bates* [2012; hereafter B12]. When this model is used, it is shown that the EfCS depends not only on the average of the two radiative response coefficients but also on the difference and product of these quantities, as well as on the coefficient of perturbation DHT. Extending the results of B12, we show here that there is a simple criterion involving the parameters of Model B that determines whether the perturbation DHT significantly influences the EfCS.

In related work, *Armour et al*. [2013] have developed a three-zone EBM, representing land, high-latitude oceans and low-latitude oceans. With the radiative response coefficient in each zone held constant, they have shown how the effective global radiative response coefficient can vary with time as the pattern of surface warming evolves in a transient warming scenario, thus providing insight into similar results found in GCMs. They did not, however, include DHT perturbations in their EBM or consider the possible influence of such perturbations on EfCS.

The conceptually oriented GCM study of *Zelinka and Hartmann* [2012] has also examined the influence of regionally varying radiative response coefficients in a warming climate. Latitudinal variations in this quantity were seen as causing increases in DHT, which in turn led to changes in the distribution of the net TOA radiation. None of the above variations, however, were pinpointed as playing a central role in EfCS, which was seen essentially as being the quotient of the globally averaged radiative forcing and the globally averaged radiative response coefficient.

One of the themes of the present paper is the additivity of feedbacks. The term feedback is used in many different senses in the climate literature (see *Stephens* [2005], *Bates* [2007], and *Roe* [2009] for general discussions). In *Bates* [2007], the following four categories were isolated for discussion:

1. F1—feedback according to the control-theory definition,
2. F2—feedback according to the electronics definition,
3. F3—stability-altering feedback as used in climate studies, and
4. F4—sensitivity-altering feedback as used in climate studies.

It was shown that these four categories are conceptually distinct and that any pair of them can be of opposite sign when applied to an appropriately chosen system. In the present paper, feedback will be used only in the F4 sense. An important assumption in the climate literature in relation to this category of feedback is that the components due to different physical processes are always additive.

The present paper extends the sensitivity results of B12, while making use of the stability results of that paper. Here unlike B12, we assume that the radiative forcing is globally uniform, but we allow the extratropical radiative response coefficient to assume different values as well as its tropical counterpart. We examine the possible influence of the perturbation DHT on EfCS for ranges of the two radiative response coefficients suggested by recent observations and transient GCM output. Also, we here go beyond the scope of B12 by investigating the additivity of sensitivity-altering feedbacks.

The plan of the paper is as follows. In section 2 the physics of the ZDM, Model A, and Model B are outlined and their expressions for EfCS are derived. In section 3 the formulism of sensitivity-altering feedback as applied to the three models is examined. In section 4 numerical EfCS values and feedback results are presented for ranges of the parameters suggested by the recent observational and GCM studies. The conclusions are presented in section 5. A list of acronyms and abbreviations is given in Appendix A, and a clarification of the relationship between the LC11 and Model A feedback formulisms is given in Appendix B.

2. The Energy Balance Models: Physical Aspects and Sensitivity Expressions

Some aspects of the physics of the ZDM, Model A, and Model B are discussed in this section; detailed derivations are given in B12. Unless otherwise stated, it is assumed that the tropical zone extends to ±30° in both models, so that the tropical and extratropical zones each occupy half the globe. Our notation, unless otherwise stated, follows that of B12. We begin by defining radiative response coefficients, which are basic to the properties of all the EBMs.

A radiative response coefficient b (units: $\mathrm{W\,m^{-2}\,K^{-1}}$) is defined as

$$b = \frac{d\mathrm{Flux}}{dT} \qquad (1)$$

Here Flux denotes net (LW + SW) upward radiative flux at TOA and T denotes surface temperature. Estimates of b for specified regions can be obtained by linearly regressing fluctuations in Flux from observations or from transient GCM output against fluctuations in T for the same regions.

When b is used without a subscript, it will be regarded as referring to the globe. Subscripts (1, 2) will be used in the context of the two-zone models to refer to the tropics and extratropics, respectively.

Reference quantities with which (b, b_1, b_2) can be compared are the equivalent blackbody radiative response coefficients (B, B_1, B_2) for the respective regions. Taking the characteristic emission temperatures for the globe, the tropics and the extratropics given by *Lindzen et al*. [2001], i.e., (254, 259, and 249) K, respectively, we find using the derivative of the Stefan-Boltzmann law that

$$(B, B_1, B_2) = (3.7,\ 3.9,\ 3.5)\ \mathrm{W\,m^{-2}K^{-1}} \qquad (2)$$

The value of B given here corresponds closely to the global LW clear-sky radiative damping rate (3.6 W m^{-2} K^{-1}) calculated by *Chung et al.* [2010] for the GCMs used in *IPCC* [2007]. This damping rate was found by calculating the global mean change in outgoing longwave radiation (OLR) per unit area in the GCMs resulting from a horizontally and vertically uniform temperature change of 1°C with all else remaining fixed. The above reference quantities do not influence the EfCS values calculated in this paper, but they enter the discussion of feedback formulisms in section 3.

2.1. The Zero-Dimensional Model

Assuming the atmosphere remains in energetic balance as the climate system responds to external forcing, the unit-area perturbation energy equation for the ZDM (obtained by dividing the global perturbation energy equation by $4\pi a^2$, a being the Earth's radius) is

$$\frac{dE'}{dt} = Q' - bT' \tag{3}$$

Here dE'/dt denotes the perturbation rate of energy absorption by the land-ocean-ice system per unit area, Q' is the unit-area external forcing, and T' is the surface temperature perturbation; in this paper, as in B12, primes denote time-dependent perturbations and Δ will denote the corresponding steady-state perturbations. Satellite measurements [e.g., *Choi et al.*, 2014] and GCM simulations [e.g., *Brown et al.*, 2014] show that an additional term may enter the right-hand side (RHS) of equation (3) as a result of random fluctuations in SW TOA radiation associated with cloud variability not caused by variations in T. Such fluctuations are here assumed not to influence the climate system's stability or its $2\times CO_2$ sensitivity (see further discussion in section 4).

The condition for the stability of the ZDM is $b > 0$. Assuming that this is satisfied, the model's steady state response to a steady forcing ΔQ is

$$\Delta T_{\mathrm{ZDM}} = \frac{\Delta Q}{b} \tag{4}$$

If ΔQ is prescribed as the radiative forcing for a CO_2 doubling, ΔT_{ZDM} is the value of EfCS provided by the ZDM; in this paper, this forcing is in all cases assigned the standard value 3.7 W m^{-2} [e.g., *IPCC*, 2007, p. 140; *IPCC*, 2013, p. 68].

2.2. The Two-Zone Models

We next consider Models A and B. These treat the tropics (zone 1) and extratropics (zone 2) as separate zones with distinct radiative response coefficients (b_1, b_2). Assuming the atmosphere remains in energetic balance in each zone, the unit-area perturbation energy equations for the zones (obtained by dividing the perturbation energy equations for the full zones by $2\pi a^2$) are

$$\frac{dE_1'}{dt} = Q' - b_1 T_1' - \widetilde{D}' \tag{5}$$

$$\frac{dE_2'}{dt} = Q' - b_2 T_2' + \widetilde{D}' \tag{6}$$

Here the notation is a direct extension of that used in the case of the ZDM, with $\widetilde{D}' = D'/2\pi a^2$, where D' denotes the perturbation DHT (atmospheric plus oceanic) from zone 1 to zone 2.

2.2.1. Model A

In Model A, the temperature perturbations in the tropical and extratropical zones are constrained to be equal; thus,

$$T_1' = T_2' \equiv T_A' \tag{7}$$

Under these circumstances, it is seen by summing equations (5) and (6) that the condition for the global stability of Model A is

$$b_1 + b_2 > 0 \tag{8}$$

Assuming that this is satisfied, the EfCS provided by Model A, obtained by taking half the sum of equations (5) and (6) for the steady state case and using equation (7), is

$$\Delta T_A = \frac{\Delta Q}{[(b_1 + b_2)/2]} \tag{9}$$

This is in close agreement with ΔT_{ZDM} as given by equation (4).

The DHT for the steady state perturbation (denoted $\Delta\widetilde{D}_A$ in the case of Model A) is obtained by taking the half the difference of equations (5) and (6) and using equations (7) and (9); thus,

$$\Delta\widetilde{D}_A = \left(\frac{b_2 - b_1}{b_1 + b_2}\right)\Delta Q \tag{10}$$

The perturbation DHT in Model A thus emerges as a diagnostic consequence and does not appear explicitly in the expression for EfCS. However, it plays an important, if implicit, role in the underlying physics of the model (see equation (20) below and the sentence following it).

Lin et al. [2005] and *Trenberth et al*. [2010] have criticized LCH01 and LC09 for the lack of an explicit perturbation DHT in their EBM, which is the equivalent of Model A. This omission is remedied in a simple way in Model B.

2.2.2. Model B

In Model B, no a priori constraint such as equation (7) is imposed on the temperature perturbations $\left(T_1', T_2'\right)$. Instead, they are taken as independent quantities and the perturbation DHT from the tropics to the extratropics (now denoted $\widetilde{D}_B'$) is explicitly parameterized as

$$\widetilde{D}_B' = d\left(T_1' - T_2'\right) \tag{11}$$

where *d* is the DHT coefficient; an estimate of its magnitude derived from observations is given below. This parameterization is consistent with the study of *Vallis and Farneti* [2009], who concluded that the meridional heat transport by the atmosphere and wind-driven ocean gyres in midlatitudes both scale with the meridional temperature gradient.

In the real climate system, the mean DHT between the tropics and extratropics plays an essential role in governing the mean temperature in both zones. Globally, it amounts to an annual mean rate of poleward energy transport (atmospheric plus oceanic) of about 10 petawatt (PW) from the tropical regions of radiative energy surplus to the extratropical regions of radiative energy deficit, with most of transport occurring in the atmosphere [e.g., *Trenberth and Caron*, 2001, Figure 7]. Assuming an extratropical emission temperature of 249 K as given by LCH01, the total OLR from the extratropics (regarded as the regions extending poleward from ±30°) is about 55 PW. Thus, the DHT from the tropics makes up almost a fifth of the annual mean heat loss to space from the extratropics.

Using observed seasonal statistics and assuming a linear dependence of the form (equation (11)) of perturbation DHT on the perturbation temperature difference between the tropics and extratropics, *Bates* [1999; see also B12, section 3.3] estimated the coefficient of perturbation DHT ($\widehat{d}$ in B12) to be approximately 1 PW K^{-1}; when divided by the area of the tropics, this corresponds to a unit-area DHT coefficient $d = 3.9$ W m^{-2} K^{-1}. This is of the same magnitude as the reference blackbody radiative response coefficients given in equation (2), suggesting that the inclusion of DHT should be considered in any zero-order energy balance model of the climate system.

Inserting equation (11) into equations (5) and (6), the perturbation energy equations for Model B become

$$\frac{dE_1'}{dt} = Q' - b_1 T_1' - d\left(T_1' - T_2'\right) \tag{12}$$

$$\frac{dE_2'}{dt} = Q' - b_2 T_2' + d\left(T_1' - T_2'\right) \tag{13}$$

The conditions for the global stability of this model are not obvious, as was the case for the ZDM and Model A. It has been shown in B12 that if the left-hand side in equations (12) and (13) are represented as ocean mixed layer terms with equal heat capacity in the two zones and the extratropical zone is taken as locally stable ($b_2 > 0$), which will always be assumed here, Model B is globally stable provided the following criterion is satisfied:

$$S_2 \equiv b_1 b_2 + d(b_1 + b_2) > 0 \tag{14}$$

Clearly, this criterion is satisfied as long as

$$b_1 + b_2\left(\frac{d}{b_2 + d}\right) > 0 \tag{15}$$

We confine our attention to cases in which equations (14) and (15) are satisfied. In this case it is clear from equation (15) that equation (8) is satisfied a fortiori; thus, in all cases of interest, Models A and B are both globally stable.

In these circumstances, it is easily seen that equations (12) and (13) give the following steady state solutions for a CO_2 doubling

$$\Delta T_1 = (b_2 + 2d)\Delta Q/S_2 \tag{16}$$

$$\Delta T_2 = (b_1 + 2d)\Delta Q/S_2 \tag{17}$$

The EfCS provided by Model B, defined as $\Delta T_B \equiv (\Delta T_1 + \Delta T_2)/2$, is then given by

$$\Delta T_B = (1 + X)\Delta T_A \tag{18}$$

where ΔT_A is defined by equation (9) and

$$X \equiv \frac{1}{S_2}\left(\frac{b_1 - b_2}{2}\right)^2 \tag{19}$$

Thus, all effects of perturbation DHT and nonuniform radiative response coefficients on the EfCS as given by Model B are encapsulated in the quantity X. It is clear that when $0 < S_2 < \infty$ (i.e., Model B stable and d finite) and $b_1 \neq b_2$, we have $X \geq 0$, showing that in all such cases $\Delta T_B \geq \Delta T_A$.

If Model B is taken as a benchmark, a criterion for the validity of the EfCS as provided by Model A, and by extension by the ZDM, is that X be small. It is seen that $\partial X/\partial d < 0$ and that $X \rightarrow 0$ for all values of (b_1, b_2) if $d \rightarrow \infty$. Thus, Model A corresponds to Model B with an infinite DHT coefficient. For finite d, it is seen that $X \rightarrow 0$ if $b_1 \rightarrow b_2$.

Under general circumstances X can be large. For example, with b_2 assigned a fixed positive value, X tends to infinity if b_1 tends to a value such that $S_2 \rightarrow 0$. In such a situation of marginal global stability, it is clear that ΔT_B becomes extremely sensitive to the values of both d and $(b_1 - b_2)$.

The steady state perturbation DHT for Model B (denoted $\Delta\widetilde{D}_B$) is obtained by substituting equations (16) and (17) into the steady state version of equation (11), whence

$$\Delta\widetilde{D}_B = \frac{d}{S_2}(b_2 - b_1)\Delta Q \tag{20}$$

It is clear that this reduces to $\Delta\widetilde{D}_A$ as given by equation (10) if $d \rightarrow \infty$. In both two-zone EBMs the perturbation DHT is directed toward the zone in which the transported energy can be more effectively radiated to space (i.e., DHT > 0 if $b_2 > b_1$ and DHT < 0 if $b_1 > b_2$). This provides physical insight into why increasing d tends to decrease the EfCS. The effect is analogous to a negative lapse rate feedback, which occurs in the context of vertical heat transfer; the feedback in that case is negative when a $2 \times CO_2$ forcing gives an increased upward flux of heat, which decreases the lapse rate and gives a more effective radiation of heat to space from the higher levels, thereby decreasing the climate sensitivity. In the case of the DHT mechanism under discussion here the sign of the sensitivity-altering effect is more definitive.

3. The Energy Balance Models: Feedback Formulisms

The formulism of sensitivity-altering feedbacks has provided a framework for evaluating the relative contribution of different climate processes to climate sensitivity. The key aspect of the formulism is that feedbacks are assumed to combine linearly [e.g., *Peixoto and Oort*, 1992, p. 26–27; *Wallace and Hobbs*, 2006, p. 445]. The formulism is extensively used in the context of the ZDM, but it has recently been adopted in the context of Model A. We shall describe its application in the above two models and examine its applicability to Model B.

3.1. The Zero-Dimensional Model

The formulism of sensitivity-altering feedbacks as applied to the ZDM can be regarded as consisting of two steps. In step 1, the expression (4) for ΔT_{ZDM} is recast in a form suggested by the feedback formula of electronics [e.g., *Smith*, 1987, equations (15.4); *Peixoto and Oort*, 1992, equation (2.10)]:

$$\Delta T_{\mathrm{ZDM}} = \frac{\Delta T_0}{1-f} \tag{21}$$

where the zero-feedback response is

$$\Delta T_0 = \Delta Q/B \tag{22}$$

and the feedback factor (not to be confused with the "feedback parameter" referred to in the Introduction) is

$$f = (B-b)/B \tag{23}$$

The zero-feedback case (in all instances chosen somewhat arbitrarily) corresponds to that where $b = B$, with positive (negative) feedback occurring when b is less than (greater than) B. (Note that *Hansen et al*. [1984] reversed the feedback/gain terminology of electronics, calling the quantity f above a gain. As noted by *Roe* [2009], this has led to some terminological confusion in the climate feedback literature. In this paper, we adhere to the electronics terminology.)

In step 2, it is assumed that Flux can regarded as a function of T and of a number of climate fields α_i (e.g., water vapor and clouds) which are themselves functions of T; thus, b can be expanded as

$$b = \frac{\partial \mathrm{Flux}}{\partial T} + \sum_i \frac{\partial \mathrm{Flux}}{\partial \alpha_i}\frac{d\alpha_i}{dT} \tag{24}$$

Identifying the first term on the RHS of equation (24) with B, we can then write equation (23) in the form $f = \sum_i f_i$, where

$$f_i = -\frac{1}{B}\frac{\partial \mathrm{Flux}}{\partial \alpha_i}\frac{d\alpha_i}{dT} \tag{25}$$

With the above assumption regarding Flux, it is seen that the feedbacks are linearly additive and the relative contributions of the different climate fields to the net value of f can be estimated using equation (25), with positive (negative) f_i tending to increase (decrease) ΔT_{ZDM}.

It is to be noted that while the above feedback formulism provides a useful means of diagnosing the relative contributions of the separate climate processes to ΔT_{ZDM}, it is only the net value of b that determines the latter quantity.

3.2. Model A

Similarly, step 1 in applying the formulism of feedbacks to Model A consists of recasting ΔT_A as given by equation (9) in the electronics feedback form

$$\Delta T_A = \frac{\Delta T_0}{1-f} \tag{26}$$

where in this case the zero-feedback response is

$$\Delta T_0 = \frac{\Delta Q}{[(B_1+B_2)/2]} \tag{27}$$

and the feedback factor is

$$f = \frac{1}{(B_1+B_2)}[(B_1-b_1)+(B_2-b_2)] \tag{28}$$

Here the zero-feedback case corresponds to that where $b_1 = B_1$ and $b_2 = B_2$. We can write equation (28) as $f = f_1 + f_2$, where $f_1 = (B_1 - b_1)/(B_1 + B_2)$ and $f_2 = (B_2 - b_2)/(B_1 + B_2)$. The feedbacks (f_1, f_2) for the two zones thus defined retain the property of combining additively to give the global feedback f.

In principle, step 2 as used in the ZDM feedback formulism could also be applied in the case of Model A by making the corresponding assumptions about (b_1, b_2) and expanding these quantities in partial derivatives as in equation (24) and using these expansions in the expressions for (f_1, f_2) above. However, the authors who have used Model A to provide EfCS values (i.e., LCH01, LC09, LC11, and B12) have not proceeded in this way.

In LCH01, expressions were developed for (b_1, b_2) in terms of surface temperature variations that allow for such an expansion, these including blackbody terms corresponding to (B_1, B_2). It was assumed in LCH01 that in the extratropics the blackbody term is the only term in the expansion. In the tropics, the terms additional to the blackbody term involved an iris effect, whereby the area of tropical cirrus normalized by a measure of cumulus coverage contracts as the surface temperature increases. In the iris effect as formulated by LCH01, the water vapor and cloud responses to surface temperature variations cannot be separated but involve a combined response to the variations in cirrus area. This combined response leads to a negative value of f_1, which significantly decreases ΔT_A.

In LC09 and LC11, the details of the iris mechanism were not explicitly invoked, but the marked difference found between the observational and GCM values of b_1 was implicitly attributed to the existence of the mechanism in the real system and its absence in the GCMs. Although it is not obvious, the feedback formulism used by LC09 and LC11 corresponds to that for Model A as expressed by equations (26), (27), and (28) above for the special case where b_2 is assumed equal to B_2. This is clarified in Appendix B, which should help to remedy some misunderstanding of LC09 and LC11 (e.g., in *IPCC* [2013, section 10.8.2.2]; it was mistakenly stated there that the simple EBM used by LC09 and LC11 was limited to the tropics).

In B12, the emphasis as far as Model A was concerned was on exploring the net effect on ΔT_A of variations in b_1, with b_2 held fixed at the value B_2. The relative contributions of the separate physical processes influencing b_1 were not examined, and no recourse was made to a feedback formulism as described above.

3.3. Model B

For a given forcing ΔQ, the expression (18) giving the EfCS for Model B is a highly nonlinear function of the model parameters (b_1, b_2, d). Unlike the case of the ZDM and Model A, it is no longer possible to write the expression for the EfCS in a form corresponding to equation (21) or (26) with a feedback factor f that is expressible as a linear combination of terms describing the effects of the different physical processes.

Step 1 in the feedback formulation could formally be taken in this case by recasting equation (18) in the electronics feedback form

$$\Delta T_B = \frac{\Delta T_0}{1 - f} \tag{29}$$

with the zero-feedback case defined, for example, as that where $(b_1, b_2, d) = (B_1, B_2, \infty)$. This gives the same expression (27) for the zero-feedback response, ΔT_0, as in the case of Model A, but the feedback factor f now takes the form

$$f = \frac{1}{1 + X}\left[\frac{(B_1 - b_1) + (B_2 - b_2)}{B_1 + B_2} + X\right] \tag{30}$$

Clearly, this expression for f cannot be written as a linear combination of terms involving (b_1, b_2, d), except in the special case where $X = 0$ (i.e., $b_1 = b_2$ or $d = \infty$), in which case f reduces to the sum of $f_1 + f_2$ as in the case of Model A.

These results suggest that if (b_1, b_2, d) are such that $X \approx 0$ does not hold, the conventional assumption that feedbacks are additive may break down. If $X \geq 1$, there is no advantage to be gained in adopting a feedback formulism, either as a means of diagnostically investigating the relative effects of different climate processes on EfCS or of arriving at the net value of EfCS. It is simpler for the latter purpose just to use the expression (18) as it stands.

4. Numerical Values and Discussion

In this section we examine the EfCS values given by Models A and B using a range of values of b_1 based on recent studies that present both observational and GCM results. Little reliable observational information is available about the value of b_2, and therefore, we consider a range of possibilities, guided by physical reasoning and the results of a GCM study that has focused on this matter. A wide range of variation of d is allowed. We examine the question of whether $X \approx 0$ in cases that are of physical interest, considering the contexts of both the real climate system and of GCMs. The answer determines whether the perturbation DHT and the interzone variation in the radiative response coefficients can have a significant influence on EfCS and also determines whether sensitivity-altering feedbacks are additive.

4.1. Estimating b_1

The difficulty of determining the global-mean radiative response coefficient b from observations is illustrated by the study of *Dessler* [2013, hereafter D13], who used monthly mean deseasonalized anomalies of TOA radiative fluxes from the Clouds and the Earth's Radiant Energy System (CERES) satellite instrument and concurrent reanalysis meteorological fields over the period of 2000 to 2010 to obtain an estimate of this quantity (in D13 the quantity corresponding to b is the negative of the quantity called the thermal damping rate). From D13's Table 1 it is seen that b is estimated to be $1.15 \pm 0.88\,\mathrm{W\,m^{-2}\,K^{-1}}$ (uncertainty 2σ). From equation (4) we then have $\Delta T_{ZDM} = 3.2$ (1.8, 13.7)°C, where the values in curly brackets correspond to the choice of the plus and minus sign in the above estimate of b, respectively. Thus, the observational study of D13 using a ZDM as the underlying EBM gives an uncertainty range of EfCS that is much wider than the *IPCC* [2013] sensitivity range.

Brown et al. [2016] have examined the relationship between anomalies in annual mean net TOA radiation and surface temperature using CERES data for the period of 2001–2014. Using the global means of the annual mean quantities, they found that the observational value of the quantity corresponding to b was $2.4\,\mathrm{W\,m^{-2}\,K^{-1}}$ (their Figure 1a). Inserting this value into equation (4) gives $\Delta T_{ZDM} = 1.5$°C, which is at the lower end of the *IPCC* [2013] range. The relationship between radiative response coefficients derived from annual mean anomalies and deseasonalized monthly mean anomalies remains a subject to be investigated. In this paper, where we wish to separate radiative and dynamical effects in the context of Models B, we work on the assumption that, as far as the tropics are concerned, values of b_1 obtained using deseasonalized monthly mean data are appropriate. *Spencer and Braswell* [1997] have pointed out that radiation in the tropics responds to daily meteorological fields, which show sharp transitions above the boundary layer between regions that are very moist and regions that are very dry. They showed that monthly mean fields retain some of these features but suggested that important information about the sharp tropical transitions is lost when averages over longer time scales are taken.

It is to be expected that, because of these sharp transitions and related features of the tropical atmosphere (e.g., high boundary layer specific humidity and dominance of cumulonimbus clouds) that distinguish it from the extratropical atmosphere (characterized by low specific humidity and the dominance of stratiform clouds), the radiative response coefficient might vary significantly between these regions. Thus, there are a priori reasons to expect that the ZDM, with a single value of the radiative response coefficient, b, could have significant limitations and that a two-zone model with distinct coefficients (b_1, b_2) could have significant advantages as a low-order conceptual model of the climate system.

Such considerations, as well as results of LCH01 on the iris mechanism, led LC09 and LC11 to use tropical data to estimate b_1. Their approach was to focus on the oceanic region of the tropical band (20°S–20°N), with the expectation that estimates of the radiative response coefficient having a narrow range of uncertainty could be obtained there. The data used were monthly mean deseasonalized anomalies from the Earth Radiation Budget Experiment (1985–1999) and CERES (2000–2008) satellite data sets, with T obtained from meteorological reanalyses. To reduce noise, time segments in which $|T'|$ exceeded a minimum of 0.1°C were selected, giving a small number of data points in (Flux′, T'). Correlations between these fluctuations were further improved by using lagged regression, where Flux′ was allowed to lag or lead T' by a number of months. A further reduction of noise was achieved by using 3-point smoothing of the data. The methodology applied to the observational data as described above was also applied to the output of 11 AMIP GCMs and 10 Coupled Model Intercomparison Project (CMIP) GCMs used in the IPCC Fourth Assessment Report (AR4) (generically CMIP3 models; see Appendix A for definitions). The AMIP and CMIP GCM outputs both extend over the period of 1985–2008.

LC11's results for the observed LW, SW, and LW + SW regression slopes for their selected region, as well as the corresponding results for the AMIP and CMIP models, are shown here in Table 1; the observed results are taken from LC11's Table 2, the AMIP results from their Table 3, and the CMIP results from their Figure 12 (see supporting information for extra detail). These results, though obtained for the restricted tropical region, were taken by LC11 as applying to the full tropical zone of the two-zone model, thus giving estimates of b_1 for the observational and the GCM cases.

Corresponding observational and GCM regression slopes from *Mauritsen and Stevens* [2015, hereafter MS15] are also presented here in Table 1. These authors used a somewhat different methodology, using data for the

Table 1. Linear Regression Slopes (Units: $\mathrm{W\,m^{-2}\,K^{-1}}$) of Anomalies in Outgoing TOA Radiation (LW, SW, and LW + SW) Against Surface Temperature in the Tropics, as Determined From Observations [$(\mathrm{Slope})_{\mathrm{obs}}$], From AMIP GCMs [$(\mathrm{Slope})_{\mathrm{AMIP}}$], and From CMIP GCMs [$(\mathrm{Slope})_{\mathrm{CMIP}}$][a]

	$(\mathrm{Slope})_{\mathrm{obs}}$	$(\mathrm{Slope})_{\mathrm{AMIP}}$	$(\mathrm{Slope})_{\mathrm{CMIP}}$
LC11, LW	5.3 ± 1.3	1.8 {−0.8, 4.4}	3.0 {0.6, 5.8}
LC11, SW	1.9 ± 2.6	−2.9 {−3.8, −0.1}	1.2 {−3.3, 3.9}
LC11, LW + SW	6.9 ± 1.8	−1.1 {−4.7, 1.0}	4.2 {0.5, 8.1}
MS15, LW	4.1 ± 0.8	2.7 {0.8, 5.4}	2.2 {0.2, 4.2}
MS15, SW	−0.9 ± 0.9	−1.4 {−4.3, 1.8}	−1.2 {−4.6, 0.8}
MS15, LW + SW	3.2 ± 1.0	1.3 {−1.1, 4.7}	1.0 {−1.1, 3.0}

[a]The uncertainty interval in the first column of figures is ±1 standard error; values in curly brackets in the other columns are the outer limits of the quantity in question. The slopes of *Lindzen and Choi* [2011; LC11] are evaluated using data for the oceanic part of the latitude band (20°S–20°N), while those of *Mauritsen and Stevens* [2015; MS15] are evaluated using data for the entire latitude band (20°S–20°N).

entire latitude band (20°S–20°N) rather than just its oceanic region. They used CERES satellite observations and Hadley Centre, Climatic Research Unit data set, version 4 observations of T (see Appendix A for definitions) for the full years 2001–2013; the observational results are taken from MS15's Table S2 (supporting information), while the AMIP GCM and CMIP GCM results, derived from models used in the IPCC AR5 report, are taken from their Table S3 (online corrected version as given at http://www.nature.com/ngeo/journal/v8/n5/extref/ngeo2414-s1.pdf). All the MS15 slopes are at lag = 0. (Note that due to a different sign convention for TOA radiative fluxes, the MS15 slopes have here been multiplied by −1 to correspond to LC11's. Note also that, apart from the CMIP5/CMIP3 difference, the distinction "CMIP5 AMIP (uncoupled)" and "CMIP5 historical" used in MS15 corresponds to the distinction AMIP/CMIP used in LC11 and here.)

An examination of the observational results in Table 1 shows the following:

1. LC11 and MS15 both find a LW regression slope whose mean exceeds the reference blackbody radiative response coefficient B_1 and whose standard error is relatively small;
2. They both find a SW regression slope whose mean is well below B_1 and whose standard error is large, making even the sign of the SW slope ambiguous;
3. The LW + SW slope is somewhat greater than the LW slope in the LC11 case and somewhat less than the LW slope in the MS15 case.

The above features in combination suggest that the LW slopes of LC11 and MS15, which are well defined and not far apart, may provide a reasonable estimate of b_1. Such a conclusion is in accordance with the results of *Cho et al*. [2012] and *Choi et al*. [2014], who showed that observational estimates of the SW slopes are dominated by the noise due to random changes in clouds not caused by SST changes, making reliable estimation of any underlying SW signal due to SST changes impossible. *Choi et al*. [2014] also showed that the LW slope in the tropics (20°S–20°N) can be reliably estimated, thus allowing reliable estimation of b_1.

The assumption that the regression slope for the latitude band 20°S–20°N is representative of the band 30°S–30°N is also more easily justified in the LW than in the SW case, in view of the generally smaller latitudinal gradient of the LW slope across the subtropics (dominated by the blackbody contribution). Clearly, the observational value of b_1 indicated here considerably exceeds the observational estimates of b discussed earlier.

An examination of the GCM results in Table 1 shows the following:

1. The results of both LC11 and MS15 show large ranges of variation of the GCM slopes and substantial disagreement between the mean GCM slopes and the mean observed slopes, for the LW, SW, and LW + SW cases.
2. In the case of LC11's results, the CMIP slopes are in closer agreement with the observed values than are the AMIP slopes; in the case of MS15's results, however, the opposite holds.

The above results of LC11 and MS15 indicate that the difference between the observed and the GCM radiative responses in the tropics is robust. The observations suggest that $b_1 \approx 5\,\mathrm{W\,m^{-2}\,K^{-1}}$, while the GCMs have values that are in general considerably smaller, have a wide spread and may extend to negative values.

The fact that both LC11's and MS15's observational findings indicate that $b_1 > B_1$ is supportive of the existence of some form of tropical iris mechanism, although these authors differ in their views of its physical basis. LC11 propose increased precipitation efficiency at warmer temperatures as the basis, while MS15 propose increased convective aggregation. The observational finding that $b_1 > B_1$ is not in accord with theoretical LW radiative calculations based on holding relative humidity fixed at various value as the temperature increases [e.g., *Pierrehumbert*, 2010, Figure 4.31]. Such calculations suggest that $d\mathrm{OLR}/dT$ should decrease significantly as T increases from extratropical to tropical values and that one should find $b_1 << B_1$. Current GCMs appear to reflect the LW radiative properties expected on the basis of such constant relative humidity calculations rather than agreeing with the tropical observations.

The observational regression slopes shown in Table 1 have been calculated using deseasonalized monthly anomalies over the relatively short period for which satellite observations of TOA radiation are available. Over this period, El Niño–Southern Oscillation is the main source of variability in tropical SST. It is assumed in this paper that the value of b_1 obtained from the LW regression slope in these circumstances is representative of the corresponding value associated with long-term forced climate change that is needed to obtain EfCS. Supporting this assumption is the high correlation found in LC11 between the LW flux and SST increments ($R \approx 0.85$ at zero lag; LC11, Figure 10); this suggests that the LW flux increments are for the most part directly linked to the SST increments and only to a minor extent determined by effects that are independent of these increments. This being the case, it is reasonable to assume that roughly the same b_1 would be associated with SST changes caused by long-term global warming. There have been some GCM comparisons in which differences between short-term and long-term values of b or its components have been found, leading to some questioning of whether short-term observational values of radiative response coefficients can be taken as representative of the long-term values [e.g., *Trenberth et al*., 2011]. Recent studies that have found such differences are *Dalton and Shell* [2013], *Dessler* [2013], and *Koumoutsaris* [2013]. However, none of these studies has carried out like-with-like comparisons providing evidence that b_1 evaluated using short-term deseasonalized LW fluxes on a monthly time scale is unrepresentative of the corresponding long-term value under forced climate change. Instead, they draw their conclusions from differences in the long-term versus short-term values of the global quantity b but without quantifying how time variations in the pattern of warming may contribute to such differences; as *Armour et al*. [2013] have shown, it is possible for b to vary in such circumstances without any variation in (b_1, b_2). *Brown et al*. [2016] have also concluded that short-term and long-term values of radiative response coefficients may differ but have done so using annual rather than monthly mean data and using only an unforced (preindustrial) GCM simulation. *Zhou et al*. [2015] have done a comparison of the cloud component of b in unforced (preindustrial) and forced ($4 \times CO_2$) CMIP5 simulations and have found small differences between the short-term and long-term values. They acknowledge that these differences are probably due to differences in the surface warming pattern on the two time scales. In contrast to the other studies cited above, they conclude that short-term observations can be used to evaluate long-term radiative response coefficients.

4.2. Estimating b_2 and d

The comments of LC09 and LC11 regarding the difficulty of estimating b_2 observationally are corroborated by the results of D13 (see his Figure S8 in the supporting information, black curve). This shows that at extratropical latitudes, the LW + SW radiative response to variations in T (global mean) is oscillatory and of ambiguous sign. LC11's physical arguments for assuming the extratropical LW + SW radiative response coefficient to be well approximated by the blackbody value ($b_2 \approx B_2$) are based on the low specific humidity and approximately unvarying 50% cloud cover in this region (see their p. 388, right column). In view of the results of *Choi et al*. [2014] showing the difficulty of obtaining reliable estimates of the SW component of the radiative response coefficient observationally (in the extratropics as well as the tropics), we consider here estimates of the LW component alone. *Pierrehumbert* [1995] has used a GCM radiation code to evaluate the clear-sky OLR as a function of low-level air temperature for various relative humidities (see his Figure 2). Choosing a base temperature characteristic of the extratropics (280 K) and the 75% RH curve, a read-off from his figure gives $d\mathrm{OLT}/dT \approx 2.1\ \mathrm{W\,m^{-2}\,K^{-1}}$. In aquaplanet experiments using two GCMs without an iris effect, *Langen and Alexeev* [2005] found an extratropical LW response coefficient of approximately $2\ \mathrm{W\,m^{-2}\,K^{-1}}$. Guided by these results, we here allow b_2 to vary in the range of (2.0, 3.5) $\mathrm{W\,m^{-2}\,K^{-1}}$.

It is assumed here that snow/ice-albedo effects, which influence a relatively small fraction of the area of the extratropical zone in our two-zone models, do not necessitate the adoption of values of b_2 below the lower limit of the above range. These effects are generally regarded as sensitivity enhancing, although their influence on a global scale is regarded as small by comparison with temperature and water vapor effects [e.g.,

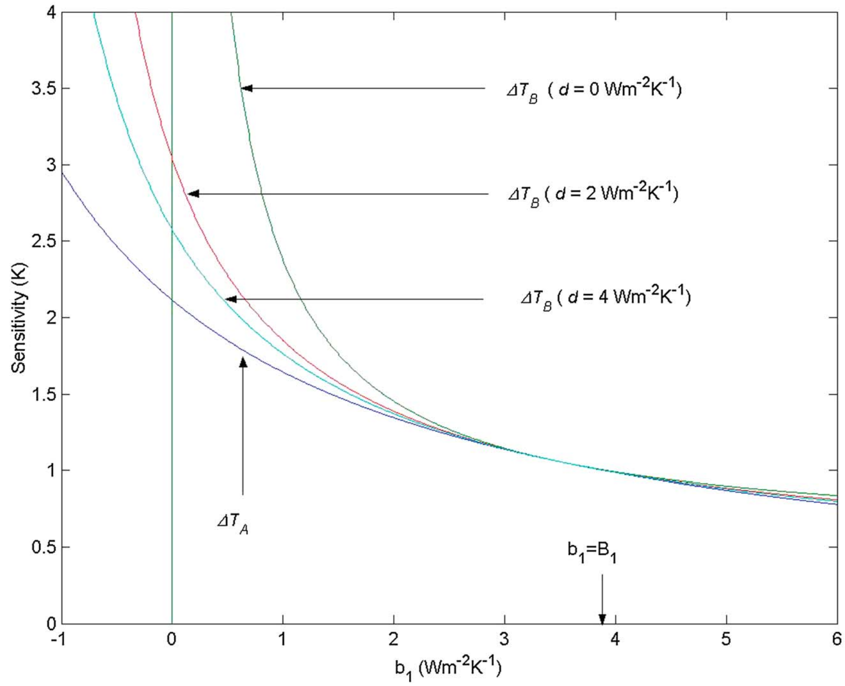

Figure 1. EfCS provided by Model A (ΔT_A) and Model B (ΔT_B) as functions of the tropical radiative response coefficient (b_1) with the extratropical radiative response coefficient (b_2) set at 3.5 W m^{-2} K^{-1} and the DHT coefficient (d) set at (0, 2, 4) W m^{-2} K^{-1}. Forcing: $\Delta Q = 3.7$ W m^{-2}. See text for further details.

Curry and Webster, 1999, section 13.5; *IPCC*, 2013, Figure 9.43]. Their influence also differs markedly between models and is not always of the same sign [e.g., *Cess et al.*, 1991; *NRC*, 2003].

We allow the DHT coefficient d in Model B to vary in the range of (0, 4) W m^{-2} K^{-1}, the larger value approximating that derived from observed seasonal statistics as discussed in section 2.2.2.

4.3. Numerical Values of ΔT_A and ΔT_B

Using LC11's LW slope from our Table 1 gives the observational estimate $b_1 = 5.3 \pm 1.3$ W m^{-2} K^{-1}, while adopting their blackbody estimate for the extratropics gives $b_2 = B_2 = 3.5$ W m^{-2} K^{-1}. Inserting these values into equation (9) then gives the observational EfCS provided by Model A as

$$(\Delta T_A)_{\text{obs}} = (0.7,\ 1.0)°\text{C} \tag{31}$$

This range of EfCS, which in round figures corresponds to that given by LC11 (see their Table 2), is narrow and lies below the lower end of the (1.5, 4.5)°C range of ECS estimates given by *IPCC* [2013]. Note that the *IPCC* [2013] ECS estimates, being based on the method of Gregory et al. [*Gregory et al.*, 2004] (see section 9.7.1 of the IPCC report), fall into the category designated EfCS in this paper.

Inserting the above estimates of (b_1, b_2) into equation (19) gives for X the range

$$(X)_{\text{obs}} = \left\{\frac{0.1}{1+0.4d},\ \frac{0.004}{1+0.5d}\right\} \tag{32}$$

This suggests that, for the real climate system, $X << 1$ for all likely values of d and therefore Model A is adequate for obtaining the EfCS. Likewise, it suggests that, for the real system, feedbacks are additive and the traditional feedbacks formulism is valid.

In the case of the GCM estimates of b_1, however, the situation is quite different. Here X can be large and the difference between $(\Delta T)_B$ and $(\Delta T)_A$ can be substantial. The relationship between these quantities for ranges of variation of (b_1, b_2, d) typical of the GCM values given in Table 1 is shown in Figures 1 and 2.

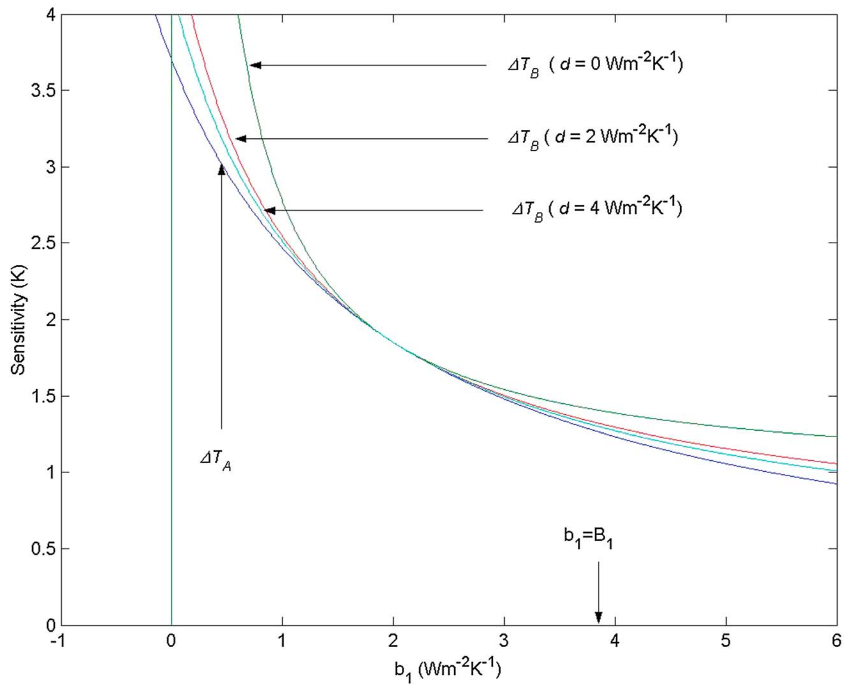

Figure 2. Same as in Figure 1 except with the extratropical radiative response coefficient (b_2) set at $2\,\mathrm{W\,m^{-2}\,K^{-1}}$.

Figure 1, with b_2 fixed at $3.5\,\mathrm{W\,m^{-2}\,K^{-1}}$, shows that for $b_1 \geq B_1$ the values of EfCS given by Models A and B are close, both being small (≤1°C) and relatively insensitive to variations in (b_1, d). For $b_1 << B_1$, on the other hand, the difference between the EfCS values can be large, becoming very sensitive to the values of (b_1, d) as b_1 tends toward zero and negative values.

Assuming $b_2 \approx 3.5\,\mathrm{W\,m^{-2}\,K^{-1}}$, LC11 have argued on the basis of Model A that, since current GCMs generally have values of b_1 much smaller than $(b_1)_{obs}$, the GCMs are considerably overestimating climate sensitivity. The results for Model B shown in Figure 1 indicate that any such underestimation of the value of b_1 by the GCMs may result in an even greater overestimation of climate sensitivity than suggested by LC11.

Figure 2 shows the results corresponding to those of Figure 1 for the case where b_2 is fixed at $2\,\mathrm{W\,m^{-2}\,K^{-1}}$. Here the values of EfCS given by Models A and B for $b_1 \geq B_1$ diverge more than in Figure 1 as b_1 increases but still remain in the neighborhood of 1°C. For $b_1 << B_1$, the difference between the EfCS values is again very sensitive to the values of (b_1, d) as b_1 tends toward zero and negative values.

The large differences between the EfCS values given by Models A and B in the regions where $b_1 << B_1$ in Figures 1 and 2 suggest that circumstances occur in present-day GCMs in which the condition $X << 1$ is far from satisfied (note that $X = (\Delta T_B - \Delta T_A)/\Delta T_A$). In these circumstances, it is clear from equation (30) that the additivity of feedbacks can break down and the traditional feedbacks formulism may no longer be useful.

The results presented in Figures 1 and 2 correspond to a broad range of the parameters (b_1, b_2, d) and are illustrative of the EfCS values for both the observational and GCM cases. In order to focus on the observational value, we present in Table 2 the range of ΔT_B corresponding to the parameter choices $b_1 = (4.1,\ 5.3)\,\mathrm{W\,m^{-2}\,K^{-1}}$ (these are the mean observational LW slopes of MS15 and LC11, respectively, from Table 1), $b_2 = (2.0,\ 3.5)\,\mathrm{W\,m^{-2}\,K^{-1}}$ (our best estimate of the likely range of this

Table 2. EfCS as Given by Model B for the Mean Observational Range of b_1 and the Best Estimates of the Likely Ranges of (b_2, d)[a]

(b_1, b_2, d)	ΔT_B
(4.1, 3.5, 2.0)	0.977
(4.1, 3.5, 4.0)	0.976
(5.3, 3.5, 2.0)	0.860
(5.3, 3.5, 4.0)	0.854
(4.1, 2.0, 2.0)	1.279
(4.1, 2.0, 4.0)	1.254
(5.3, 2.0, 2.0)	1.123
(5.3, 2.0, 4.0)	1.083

[a]Units of (b_1, b_2, d): $\mathrm{W\,m^{-2}\,K^{-1}}$. Units of ΔT_B: °C. Forcing: $\Delta Q = 3.7\,\mathrm{W\,m^{-2}}$. See text for further details.

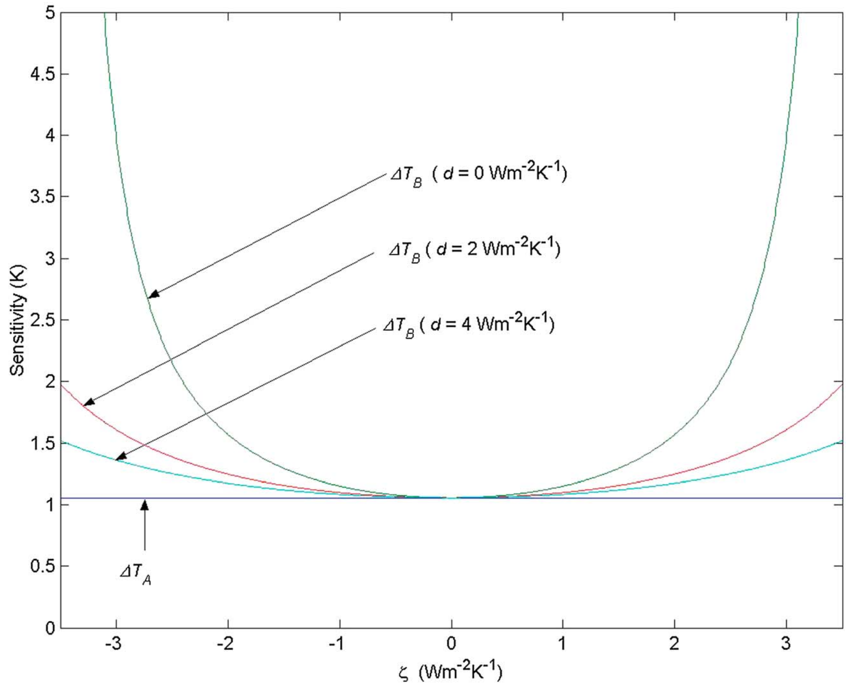

Figure 3. EfCS as provided by Model A (ΔT_A) and Model B (ΔT_B) as functions of ζ (W m^{-2} K^{-1}), where the tropical and extratropical radiative response coefficients are given by $b_1 = b + \zeta$, $b_2 = b - \zeta$ ($b = 3.5$ W m^{-2} K^{-1}) and the DHT coefficient (d) is set at (0, 2, 4) W m^{-2} K^{-1}. Forcing: $\Delta Q = 3.7$ W m^{-2}. See text for further details.

parameter), and $d = (2, 4)$ W m^{-2} K^{-1} (our central values of this parameter, regarding the values 0 and ∞ as less likely). The numbers in the first four rows of Table 2 are illustrative of the observational region of Figure 1, those in the second four rows of the observational region of Figure 2 .

It can be seen that the values of ΔT_B in the table lie in the neighborhood of 1°C and their range of variation is small (mean value 1.051°C, range {0.854, 1.279}°C). Thus, our best EfCS/Model B estimates lies below the current IPCC climate sensitivity range of (1.5, 4.5)°C and are much more tightly constrained.

In Figure 3 we show the results for the case where the sum $(b_1 + b_2)/2$ remains fixed at b, with b set to 3.5 W m^{-2} K^{-1}, while $b_1 = b + \zeta$, $b_2 = b - \zeta$, ζ being a parameter varying in the range of ± 3 W m^{-2} K^{-1}. In this case the EfCS given by Model A (and the ZDM) remains fixed at 1.06°C, but the EfCS given by Model B can vary by large amounts as the parameters (b_1, b_2, d) change.

This figure provides insight into the sensitivity of EfCS that is additional to the considerations of LC11 (see their Figure 11 and related text) [see also *Roe and Baker*, 2007]. There, the primary quantity to which the EfCS is regarded as sensitive is the feedback factor f, i.e., referring to Appendix B, equation (B8) below, the quantity $(B_1 - b_1)/(B_1 + B_2)$. The results shown in Figure 3 suggest that, in circumstances prevailing in current GCMs, the EfCS may also be very sensitive to the quantities $(b_1 - b_2)$ and d.

The combined results of Figures 1–3 suggest that, in order to reduce the large spread of ECS estimates given by current GCMs, the deficiencies in the GCM representations of b_1 pointed out by LC11 and MS15 should be rectified as a matter of priority. At present, most of the effort in tuning GCMs is focused on satisfying global mean TOA energy balance constraints [e.g., *IPCC*, 2013, Box 9.1]. Successful tuning in this regard does not ensure that the separate values of (b_1, b_2, d) are well represented. Our Model B results indicate the importance of achieving a realistic representation of all three of these parameters. Ultimately, satellite observations may provide a means of achieving better GCM representation not only of these quantities, which are indicated as being of the highest importance by a simple two-zone EBM, but also of the quantities determining the more general spatial variability of climate variables seen in climate change studies [e.g., *Ma and Xie* 2013, and references therein].

4.4. EfCS/Two-Zone EBM Values Poorly Correlated With the Standard ECS Estimates in GCMs

General acceptance of the low observational EfCS/two-zone EBM values discussed here has been inhibited by the poor correlations found in the GCM context between the sensitivity values given by the EfCS/two-zone EBM method and the standard GCM sensitivity estimates. However, the two-zone EBM used in the GCM comparisons was Model A, which we have shown to be inappropriate since many of the GCMs have $b_1 << B_1$.

Thus, LC11 found such low correlations using the output of 11 GCMs from IPCC AR4 (see their Table 4). Column 2 in this table gives the EqCS values, while column 3 gives the corresponding EfCS/Model A values. Seen explicitly in terms of Model A, the EfCS values were obtained using our equation (9) with b_1 as given by LC11's LW + SW values of $(\text{Slope})_{\text{AMIP}}$ summarized in our Table 1 and with b_2 assigned the fixed value of 3.3 W m^{-2} K^{-1}. The poor correlation between the sensitivities in columns 2 and 3 is apparent by inspection and has been pointed out in quantitative terms by MS15; they found that, omitting one model with infinite EfCS, the actual correlation was −0.11, although LC11 pointed out that all the EqCS values lay within the 90% confidence intervals of their EfCS values.

MS15 have done a similar comparison using ECS estimates given by the IPCC AR5 GCMs (listed in the right-hand column of their Table S3); however, instead of comparing these ECS estimates directly with the EfCS/Model A values, in the manner of LC11, they examined the correlation between the inverse of the ECS estimates and their LW + SW values of $(\text{Slope})_{\text{AMIP}}$ summarized in our Table 1. For this case they found a correlation of +0.38. Proceeding similarly using their $(\text{Slope})_{\text{CMIP}}$ values summarized in our Table 1 they found a correlation of +0.32. (Note that these two correlation values are taken from the corrected online version of MS15, published at http://www.nature.com/ngeo/journal/v8/n5/full/ngeo2414.html#correction1. They replace the values of +0.42 and +0.15, respectively, given on p. 347 in the original version.) Although the MS15 correlations are not as weak as those of LC11, they can still be interpreted as indicating that low confidence should be placed in the EfCS/two-zone EBM method.

However, a true GCM comparison between the EfCS/two-zone EBM method and the standard methods of estimating ECS would require a GCM correlation study in which the two-zone EBM was Model B and in which all three of the parameters (b_1, b_2, d) were evaluated from the GCMs themselves. Such a study would constitute a major research undertaking. On current evidence, the poor correlations found in the GCM comparisons that use Model A as the basis of the EfCS/two-zone EBM method, and that evaluate only b_1 from the GCMs themselves, cannot be seen as providing an argument against the EfCS/two-zone EBM method and the low observational sensitivity estimates it provides.

5. Conclusions

Some aspects of estimating $2 \times CO_2$ equilibrium climate sensitivity using observations and transient GCM output have been examined. Sensitivity values obtained using this approach are known as effective climate sensitivity estimates. All such estimates involve the use of an accompanying energy balance model. Of these, the zero-dimensional model, in which all quantities are global and annual means, has long occupied the dominant position in the climate literature. The question of whether this model is adequate for estimating effective climate sensitivity in all circumstances has been examined and the possible advantages of using two-zone (tropical/extratropical) energy balance models have been considered.

The use of two-zone models has been motivated by both theoretical and observational considerations. On the theoretical side, the different physical characteristics of the tropical and extratropical atmospheres suggest that the values of the radiative response coefficient for these regions might differ substantially and that simply using a global mean coefficient might not be sufficient for estimating effective climate sensitivity. On the observational side, the difficulty of obtaining well-defined values of the global-mean radiative response coefficient has led to a search for better defined local values in the tropics. Recent observational studies have shown that, because of random internal fluctuations in cloud cover not caused by surface temperature variations, the SW component of the radiative response coefficient cannot be determined reliably, either globally or in the tropics. However, these studies have shown that the LW component in the tropics can be determined reliably, that it generally exceeds the reference blackbody radiative response coefficient for the tropics (3.9 W m^{-2} K^{-1}), and that it considerably exceeds estimates of its global mean counterpart.

Concurrently, studies of GCM output have shown that the tropical radiative response coefficient varies widely between GCMs and that there are robust differences between the GCM and observational values of this quantity. In particular, the LW GCM values in the tropics are, in general, considerably smaller than the LW observational values and are even negative in some instances.

It is of significance for the conclusions of this paper that the findings of *Lindzen and Choi* [2011] to the above effect have recently been corroborated by the independent study of *Mauritsen and Stevens* [2015]. The latter study has also made it clear that the lack of agreement between the GCM and observational LW radiative responses in the tropics can be even more marked when the GCMs are run with a coupled ocean than without. Here we have studied the influence on effective climate sensitivity estimates of inserting the observational and GCM results of both of these studies into two distinct two-zone energy balance models, termed Models A and B. Model A, which has no explicit parameterization of perturbations in dynamical heat transport between the tropical and extratropical zones, has been criticized for not including this effect. This omission has been remedied in a simple way in Model B by inserting an explicit parameterization of the perturbation dynamical heat transport that depends linearly on the difference between the tropical and extratropical temperature perturbations.

Assuming the radiative forcing due to a CO_2 doubling is globally uniform (its value is here taken as 3.7 W m^{-2}), the effective climate sensitivity estimate given by Model A depends only on the mean of the tropical and extratropical radiative response coefficients. The corresponding estimate given by Model B, however, depends in a complex way on the separate values of these coefficients and on the dynamical heat transport coefficient. In this paper, all three of these parameters have been allowed to vary within limits suggested by the available observational, theoretical, and GCM evidence, and the resulting ranges of effective climate sensitivity have been examined. The results can be summarized as follows.

1. Using the observational estimates of the tropical radiative response coefficient ($\approx$ 5 W m^{-2} K^{-1}) and allowing the extratropical coefficient to assume values in its probable range ({2.0, 3.5} W m^{-2} K^{-1}), the estimates of effective climate sensitivity given by Models A and B are found to be close and quite insensitive to the value of the dynamical heat transport coefficient. The estimates given by both models in this case lie in the neighborhood of 1°C.
2. These estimates are considerably smaller and more tightly constrained than recent estimates obtained using a zero-dimensional model with a single global-mean value of the radiative response coefficient determined from global-mean observations. It has been argued here that the two-zone approach to making observational estimates is to be preferred.
3. Using the GCM values of the tropical radiative response coefficient and again allowing its extratropical counterpart to vary in the range (2.0, 3.5) W m^{-2} K^{-1}, the estimates of effective climate sensitivity given by Models A and B are found to be generally much larger than 1°C, and the estimates given by Model B can substantially exceed those given by Model A. In these circumstances, the estimates given by Model B can depend sensitively on the values of all three of its parameters. In particular, any underestimation of the value of the tropical radiative response coefficient by comparison with observations, such as has been indicated by recent studies to exist in current GCMs, can cause the GCMs to give a substantial overestimation of the effective climate sensitivity.
4. Related to the above, comparisons in the GCM context between climate sensitivity estimates obtained using the standard methods and effective climate sensitivity estimates obtained using a two-zone energy balance model have found a poor correlation between the two sets of estimates. This has induced a lack of confidence in the two-zone method and a lack of acceptance of the low observational climate sensitivity estimates it provides. It has been pointed out here, however, that this argument against the two-zone method and its low observational estimates lacks validity, since the two-zone model used in the comparison was Model A, which is inappropriate in the GCM context.
5. It has been shown that the additivity property of sensitivity-altering feedbacks generally assumed in the climate literature is valid for Model A and is supported by the results of Model B in the case where the observational values of the tropical radiative response coefficient are used. However, it is not generally supported by Model B in cases where the GCM values are used. A well-defined criterion for the validity of the additivity property in terms of the three parameters of Model B has been provided.

6. The demonstration using Model B that climate sensitivity may depend not only on the mean of the tropical and extratropical radiative response coefficients but also on the difference between them suggests that interzone differences in radiative responses should receive increased focus in GCM tuning efforts. Such efforts currently focus mainly on global means.

The central conclusion of this study is that to disregard the low values of effective climate sensitivity (≈1°C) given by observations on the grounds that they do not agree with the larger values of equilibrium, or effective, climate sensitivity given by GCMs, while the GCMs themselves do not properly represent the observed value of the tropical radiative response coefficient, is a standpoint that needs to be reconsidered.

Appendix A: Acronyms and Abbreviations

AMIP	Atmospheric Model Intercomparison Project
AMIP GCM	A GCM with atmosphere and land surface only, using observed SST and sea ice extent, as used in the AMIP
AOGCM	Atmosphere-ocean GCM (includes the cases with a dynamic ocean and a mixed layer ocean)
CERES	Clouds and the Earth's Radiant Energy System (instrument)
CMIP	Coupled Model Intercomparison Project
CMIP GCM	A GCM run in coupled atmosphere-ocean mode, as used in the CMIP
CMIP3	CMIP phase 3
CMIP5	CMIP phase 5
DHT	dynamical heat transport (atmospheric plus oceanic, from the tropics to the extratropics)
EBM	energy balance model
ECS	equilibrium climate sensitivity (of the real climate system)
EfCS	effective climate sensitivity (estimate of ECS obtained using observations or transient GCM output and an accompanying EBM)
EqCS	equilibrium climate sensitivity (estimate of ECS obtained by running an AOGCM to equilibrium)
ERBE	Earth Radiation Budget Experiment
GCM	global climate model (includes AMIP GCM, CMIP GCM, and AOGCM)
HadCRUT4	Hadley Centre, Climatic Research Unit (University of East Anglia) data set, version 4
IPCC	Intergovernmental Panel on Climate Change
IPCC AR4	IPCC Fourth Assessment Report (2007)
IPCC AR5	IPCC Fifth Assessment Report (2013)
LW	longwave
OLR	outgoing longwave radiation
PW	petawatt (10^{15} W)
SST	sea surface temperature
SW	shortwave
TOA	top of the atmosphere
ZDM	zero-dimensional model

Appendix B: Correspondence Between the LC11 and Model A Feedback Formulisms

In the version of Model A underlying the feedback formulism of LC09 and LC11, the areas of the tropical and extratropical zones were allowed to be unequal. We generalize our Model A formulism here to cover that case by assuming that these zones occupy fractions (γ, $1-\gamma$) of the globe, respectively. For clarity, we also here allow different time-invariant radiative forcings ($\Delta Q_1, \Delta Q_2$) in the two zones.

Using the assumption (7), it is easily seen that the perturbation energy equations (5) and (6) then generalize to

$$2\gamma \frac{dE_1'}{dt} = 2\gamma\left(\Delta Q_1 - b_1 T_A'\right) - \widetilde{D}_A' \tag{B1}$$

$$2(1-\gamma)\frac{dE_2'}{dt} = 2(1-\gamma)\left(\Delta Q_2 - b_2 T_A'\right) + \widetilde{D}_A' \tag{B2}$$

Adding these equations and assuming a steady state, we find

$$\Delta T_A = \frac{\gamma \Delta Q_1 + (1-\gamma)\Delta Q_2}{\gamma b_1 + (1-\gamma) b_2} \tag{B3}$$

Defining a globally averaged blackbody radiative response coefficient as

$$B_A = \gamma B_1 + (1-\gamma) B_2$$

it is seen that equation (B3) can be written in the feedback form

$$\Delta T_A = \frac{\Delta T_0}{1-f} \tag{B4}$$

where

$$\Delta T_0 = [\gamma \Delta Q_1 + (1-\gamma)\Delta Q_2]/B_A \tag{B5}$$

$$f = -[\gamma(b_1 - B_1) + (1-\gamma)(b_2 - B_2)]/B_A \tag{B6}$$

The above formulation can be identified with that of LC11 if it is assumed, as in LC11, that the forcing is globally uniform ($\Delta Q_1 = \Delta Q_2 = \Delta Q$) and that $b_2 = B_2$. In that case, equations (B5) and (B6) become

$$\Delta T_0 = \Delta Q / B_A \tag{B7}$$

and

$$f = -\gamma(b_1 - B_1)/B_A \tag{B8}$$

Clearly, equation (B7) corresponds to LC11's equation (1) with B_A replaced by the equivalent quantity $1/G_0$. Also, identifying B_1 here with LC11's $(\mathrm{ZFB}/\Delta\mathrm{SST})_{\mathrm{tropics}}$, it is seen that f as given by equation (B8) corresponds to LC11's feedback factor f as defined by their equation (6) provided we take $\gamma = 1/c$; thus, LC11's factor c is identified as the inverse of the fraction of the globe occupied by the tropical zone in their two-zone model. The standard value used by LC11 was $c = 2$ (i.e., $\gamma = 0.5$), which corresponds to allowing the tropical zone to occupy half the globe.

Acknowledgments
I thank Yong-Sang Choi for providing the data shown in Table S1 in the supporting information; any additional details may be obtained from Choi (e-mail: ysc@ewha.ac.kr). I also thank Thorsten Mauritsen for providing the advance information on corrections to data in the *Mauritsen and Stevens* [2015] paper; these corrections have now been published online at http://www.nature.com/ngeo/journal/v8/n5/full/ngeo2414.html#correction1 and http://www.nature.com/ngeo/journal/v8/n5/extref/ngeo2414-s1.pdf. The data provided by these authors have been very helpful in constructing Table 1. I am also grateful to the reviewers, whose comments have led to significant improvements in the manuscript.

References

Armour, K. C., C. M. Bitz, and G. H. Roe (2013), Time-varying climate sensitivity from regional feedbacks, *J. Clim.*, *26*, 4518–4534, doi:10.1175/JCLI-D-12-00544.1.
Bates, J. R. (1999), A dynamical stabilizer in the climate system: A mechanism suggested by a simple model, *Tellus*, *51A*, 349–372.
Bates, J. R. (2007), Some considerations of the concept of climate feedback, *Q. J. R. Meteorol. Soc.*, *133*, 545–560 [Erratum: p. 1071].
Bates, J. R. (2012), Climate stability and sensitivity in some simple conceptual models, *Clim. Dyn.*, *38*, 455–473, doi:10.1007/s00382-010-0966-0.
Brown, P. T., W. Li, L. Li, and Y. Ming (2014), Top-of-atmosphere radiative contribution to unforced decadal global temperature variability in climate models, *Geophys. Res. Lett.*, *41*, 5175–5183, doi:10.1002/2014GL060625.
Brown, P. T., W. Li, J. H. Jiang, and H. Su (2016), Unforced surface air temperature variability and its contrasting relationship with the anomalous TOA energy flux at local and global spatial scales, *J. Clim.*, *29*, 925–940.
Cess, R. D., et al. (1991), Interpretation of snow-climate feedback as produced by 17 general circulation models, *Science*, *253*, 888–892.
Charney Report (1979), Carbon dioxide and climate: A scientific assessment, National Academy of Sciences, Washington, D. C.
Cho, H., C. H. Ho, and Y. S. Choi (2012), The observed variation in cloud-induced longwave radiation in response to sea surface temperature over the Pacific warm pool from MTSAT-1R imagery, *Geophys. Res. Lett.*, *39*, L18802, doi:10.1029/2012GL052700.
Choi, Y.-S., H. Cho, C.-H. Ho, R. S. Lindzen, S.-K. Park, and X. Yu (2014), Influence of non-feedback variations of radiation on the determination of climate feedback, *Theor. Appl. Climatol.*, *115*, 355–364.
Chung, E.-S., D. Yeomans, and B. J. Soden (2010), An assessment of climate feedback processes using satellite observations of clear-sky OLR, *Geophys. Res. Lett.*, *37*, L02702, doi:10.1029/2009GL041889.
Curry, J. A., and P. J. Webster (1999), *Thermodynamics of Atmospheres and Oceans*, Academic Press.
Dalton, M. M., and K. M. Shell (2013), Comparison of short-term and long-term radiative feedbacks and variability in twentieth-century global climate model simulations, *J. Clim.*, *26*, 10051–10070.
Dessler, A. E. (2013), Observations of climate feedbacks over 2000-10 and comparisons to climate models, *J. Clim.*, *26*, 333–342, doi:10.1175/JCLI-D-11-00640.1.
Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative forcing and climate sensitivity, *Geophys. Res. Lett.*, *31*, L03205, doi:10.1029/2003GL018747.
Hansen, J., A. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lerner (1984), Climate sensitivity: Analysis of feedback mechanisms, in *Climate Processes and Climate Sensitivity*, *AGU Geophys. Monogr., 29*, vol. 5, edited by M. Ewing, J. E. Hansen, and T. Takahashi, pp. 130–163, AGU, Washington, D. C.
Intergovernmental Panel on Climate Change (IPCC) (2007), Climate Change 2007: The Physical Science Basis, in *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K., and New York.

Intergovernmental Panel on Climate Change (IPCC) (2013), Climate Change 2013: The Physical Science Basis, in *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climatse Change*, edited by T. F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K., and New York.

Koumoutsaris, S. (2013), What can we learn about climate feedbacks from short-term climate variations?, *Tellus A*, *65*, 18887, doi:10.3402/tellusa.v65i0.18887.

Langen, P. L., and V. A. Alexeev (2005), A study of non-local effects of Budyko-Sellers infrared parameterizations using atmospheric general circulation models, *Tellus A*, *57*, 654–661.

Lin, B., T. Wong, B. A. Wielicki, and Y.-X Hu (2005), Reply, *J. Clim.*, *18*, 2128–2131, doi:10.1175/JCLI3393.1.

Lindzen, R. S., and Y.-S. Choi (2009), On the determination of climate feedbacks from ERBE data, *Geophys. Res. Lett.*, *36*, L16705, doi:10.1029/2009GL039628.

Lindzen, R. S., and Y.-S. Choi (2011), On the observational determination of climate sensitivity and its implications, *Asia-Pacific J. Atmos. Sci.*, *47*, 377–390.

Lindzen, R. S., M.-D. Chou, and A. Y. Hou (2001), Does the Earth have an adaptive infrared iris?, *Bull. Am. Meteorol. Soc.*, *82*, 417–432.

Ma, J., and S.-P. Xie (2013), Regional patterns of sea surface temperature change: A source of uncertainty in future projections of precipitation and atmospheric circulation, *J. Clim.*, *26*, 2482–2501.

Mauritsen, T., and B. Stevens (2015), Missing iris effect as a possible cause of muted hydrological change and high climate sensitivity in models, *Nat. Geosci.*, *8*, 346–351, doi:10.1038/ngeo2414.

NRC (2003), Estimating climate sensitivity, Report of a Workshop, National Research Council, Washington, D. C.

Peixoto, J. P., and A. H. Oort (1992), *Physics of Climate*, 520 pp., American Institute of Physics, New York.

Pierrehumbert, R. T. (1995), Thermostats, radiator fins and the local runaway greenhouse, *J. Atmos. Sci.*, *52*, 1784–1806.

Pierrehumbert, R. T. (2010), *Principles of Planetary Climate*, 652 pp., Cambridge Univ. Press.

Roe, G. (2009), Feedbacks, timescales and seeing red, *Ann. Rev. Earth. Planet. Sci.*, *37*, 93–115.

Roe, G. H., and M. B. Baker (2007), Why is climate sensitivity so unpredictable?, *Science*, *318*, 629–632, doi:10.1126/science.1144735.

Smith, R. J. (1987), *Electronics: Circuits and Devices*, 3rd ed., John Wiley.

Spencer, R. W., and W. D. Braswell (1997), How dry is the tropical free troposphere? Implications for global warming theory, *Bull. Am. Meteorol. Soc.*, *78*, 1097–1106.

Stephens, G. L. (2005), Cloud feedbacks in the climate system: A critical review, *J. Clim.*, *18*, 237–273.

Trenberth, K. E., and J. M. Caron (2001), Estimates of meridional atmosphere and ocean heat transports, *J. Clim.*, *14*, 3433–3443.

Trenberth, K. E., J. T. Fasullo, T. O'Dell, and T. Wong (2010), Relationships between tropical sea surface temperature and top-of-atmosphere radiation, *Geophys. Res. Lett.*, *37*, L03702, doi:10.1029/2009GL042314.

Trenberth, K. E., J. T. Fasullo, and J. P. Abraham (2011), Issues in establishing climate sensitivity in recent studies, *Remote Sens.*, *3*, 2051–2056, doi:10.3390/rs3092051.

Vallis, G. K., and F. Farneti (2009), Meridional energy transport in the coupled atmosphere-ocean system: Scaling and numerical experiments, *Q. J. R. Meteorol. Soc.*, *135*, 1643–1660.

Wallace, J. M., and P. V. Hobbs (2006), *Atmospheric Science: An Introductory Survey*, 2nd ed., 483 pp., Elsevier/Academic Press.

Zelinka, M. D., and D. L. Hartmann (2012), Climate feedbacks and their implications for poleward energy flux changes in a warming climate, *J. Clim.*, *25*, 608–624, doi:10.1175/JCLI-D-11-00096.1.

Zhou, C., M. D. Zelinka, A. E. Dessler, and S. A. Klein (2015), The relationship between interannual and long-term cloud feedbacks, *Geophys. Res. Lett.*, 10,463–10,469, doi:10.1002/2015GL066698.