# Quantifying the Sources of Intermodel Spread in Equilibrium Climate Sensitivity

PETER M. CALDWELL, MARK D. ZELINKA, KARL E. TAYLOR, AND KATE MARVEL*

*Lawrence Livermore National Lab, Livermore, California*

(Manuscript received 15 May 2015, in final form 1 October 2015)

ABSTRACT

This study clarifies the causes of intermodel differences in the global-average temperature response to doubled $CO_2$, commonly known as equilibrium climate sensitivity (ECS). The authors begin by noting several issues with the standard approach for decomposing ECS into a sum of forcing and feedback terms. This leads to a derivation of an alternative method based on linearizing the effect of the net feedback. Consistent with previous studies, the new method identifies shortwave cloud feedback as the dominant source of intermodel spread in ECS. This new approach also reveals that covariances between cloud feedback and forcing, between lapse rate and longwave cloud feedbacks, and between albedo and shortwave cloud feedbacks play an important and previously underappreciated role in determining model differences in ECS. Defining feedbacks based on fixed relative rather than specific humidity (as suggested by Held and Shell) reduces the covariances between processes and leads to more straightforward interpretations of results.

## 1. Introduction

Equilibrium climate sensitivity (ECS; the global-average equilibrium surface temperature change due to $CO_2$ doubling) is one of our primary measures of the severity of anthropogenic climate change. Because global warming is the result of complex interactions between many aspects of the climate system, estimates of ECS are typically made using sophisticated general circulation models (GCMs). Discrepancies between ECS predictions from different GCMs are often taken as a proxy for global warming uncertainty (Flato et al. 2013), even though intermodel differences do not account for biases shared across models and therefore actually constitute a lower bound on the true uncertainty.

To reduce intermodel spread and thus (hopefully) reduce uncertainty about climate change, many researchers have focused on identifying the root cause of intermodel differences in ECS. The starting point for these efforts is typically the equation describing the response of the top-of-atmosphere energy budget to a radiative forcing $F$:

$$\Delta N = F + \lambda \Delta T. \qquad (1)$$

In Eq. (1), $F$ induces a top-of-atmosphere (TOA) energy imbalance $\Delta N$ that is partially mitigated by climate feedbacks involving a change in global-averaged surface temperature $\Delta T$ modulated by the net feedback strength $\lambda$. In equilibrium, the planet neither gains nor loses heat, so $\Delta N = 0$. By definition, in equilibrium $\Delta T$ = ECS, so

$$\mathrm{ECS} = \frac{-F}{\lambda}. \qquad (2)$$

Historically, $F$ was taken as the direct radiative impact of $CO_2$ doubling after accounting for rapid stratospheric adjustment, but Gregory et al. (2004) noted that the troposphere also experiences rapid adjustments related to the direct (rather than temperature-mediated) impact of increased $CO_2$. Here (as in most modern studies), rapid adjustments and direct $CO_2$ effect are combined to form an "effective forcing."

The net feedback $\lambda$ can be broken down into a sum of individual feedback mechanisms $\lambda_i$. The exact decomposition into individual feedbacks varies somewhat between studies, but the most common partitioning differentiates between Planck (Pl), water vapor (WV), lapse rate (LR), albedo (Alb), and cloud (Cld) terms.

* Current affiliation: NASA Goddard Institute for Space Studies, New York, New York.

*Corresponding author address*: Peter M. Caldwell, L-103, Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94566.
E-mail: caldwell19@llnl.gov

DOI: 10.1175/JCLI-D-15-0352.1

The feedback term $\lambda_{Pl}$ represents the rapid increase in longwave emission with temperature as encapsulated by Planck's law; it is negative in sign (acting to stabilize the climate) and has the largest amplitude of all feedbacks. Because water vapor increases with warming and acts as an additional greenhouse gas, $\lambda_{WV}$ is positive. The feedback term $\lambda_{LR}$ is negative because air higher in the troposphere warms more than air near the surface as the planet warms, allowing the planet to radiate to space more efficiently as surface temperature increases. The feedback term $\lambda_{Alb}$ tends to be small and positive and is related to decreases in future snow and/or ice coverage and changes in solar reflectance due to vegetation changes. Clouds have complex effects on TOA radiation. In the shortwave (SW), clouds act primarily to cool the planet by reflecting solar radiation back to space. In the longwave (LW), clouds warm the planet by trapping infrared radiation. SW cloud radiative effects depend primarily on incoming solar radiation (which varies with latitude and season), cloud fraction, and cloud albedo. LW cloud radiative effects depend primarily on cloud fraction and cloud-top altitude. Cloud feedback is positive in most models because cloud amount tends to decrease and cloud-top altitude tends to increase as the planet warms (Zelinka et al. 2012, 2013). These positive feedback effects are slightly offset by robust increases in high-latitude cloud optical depth (Gordon and Klein 2014). Although these behaviors are captured by most GCMs, the magnitude of global warming–induced changes in cloud properties differs between models in subtle but important ways, which lead to large intermodel spread in the magnitude of cloud feedback.

Several methods have been proposed for identifying the processes responsible for intermodel spread in ECS. One approach is to simply compare intermodel variance in each feedback mechanism $\lambda_i$. Since $\lambda_i$ has units of watts per meter squared per kelvin, this is fundamentally a comparison of TOA radiative perturbations normalized by the total amount of warming in each model. Variance across models (var) is found to be much larger for $\lambda_{Cld}$ than for other feedback mechanisms. Contributions from var($\lambda_{WV}$) and var($\lambda_{LR}$) are also sizable but largely cancel when combined (e.g., Bony et al. 2006; Soden and Held 2006; Huybers 2010). Using $\lambda$ as a proxy for ECS is an incomplete solution because it neglects $F$ and because ECS is not proportional to $\lambda$. Another approach that has been used to identify sources of ECS spread is to compute intermodel spread in ECS [using Eq. (2)] with all but one term replaced by multimodel mean values. Using this method, Webb et al. (2013) showed intermodel spread in net feedback to be twice as important to ECS as forcing, with cloud feedback again being the dominant source of spread. This approach is also somewhat unsatisfactory because the variances associated with individual processes do not sum to the total variance in ECS. On a related note, Dufresne and Bony (2008, hereafter DB08) point out that the spread associated with a process of interest could equally be well defined as the decrease in var(ECS) due to replacing that process with its multimodel mean while allowing all other processes to vary across models. This new definition would provide a different estimate of the importance of that term to var(ECS), leading DB08 to conclude that approaches involving replacing some terms with multimodel values yield nonunique solutions. To avoid this arbitrariness, DB08 propose a third approach based on decomposing each model's ECS into temperature perturbations associated with each feedback and forcing mechanism. A major goal of this paper is to show that the DB08 approach is equally arbitrary.

Another aim of this study is to clarify the role of covariances between feedbacks or between forcing and feedback terms in understanding intermodel spread in ECS. Previous studies have identified correlations between $\lambda_{WV}$ and $\lambda_{LR}$ [as noted above and discussed in Cess (1975); Held and Soden 2000; Soden and Held 2006], between $\lambda_{Cld}$ and $\lambda_{Alb}$ (Huybers 2010; Mauritsen et al. 2013), between $\lambda_{Cld}$ and $\lambda_{WV+LR}$ (Huybers 2010; Mauritsen et al. 2013), and between $F$ and $\lambda_{SW\_Cld}$ (Andrews et al. 2012; Webb et al. 2013; Ringer et al. 2014). These previous studies point out in a general way that feedback and forcing terms are not independent. Our study provides a way to quantify the relative importance of spread in individual forcing or feedback terms versus covariance between those terms.

We begin in section 2 by describing how we obtain estimates of ECS, forcing, and feedback terms. In section 3 we describe the DB08 approach to partitioning ECS into a sum of terms and highlight some problems with their methodology. An alternative approach that avoids these problems is presented in section 4. In section 5 we provide a formal method for breaking the variance of a sum into a sum of covariances and apply this decomposition to both partitioning approaches. Resulting insights into the source of ECS spread are discussed in section 6, and conclusions are drawn in section 7.

## 2. Data

Following Eq. (1), we compute forcing $F$ and net feedback $\lambda$ for each model as the $y$ intercept and slope (respectively) of the linear regression of $\Delta N$ onto $\Delta T$. Once $F$ and $\lambda$ are computed, ECS is calculated following Eq. (2). This approach, which was pioneered by Gregory et al. (2004), is based on the realization that while $\Delta N$

and $\Delta T$ respond nonlinearly in time to an abrupt $CO_2$ increase, their relationship tends to be approximately linear. The climate-change signal $\Delta$ is computed as the difference between yearly samples from 150 years of abrupt quadrupling of atmospheric $CO_2$ experiment (abrupt4$\times CO_2$) data and contemporaneous 21-yr running-mean values from piControl runs. The choice to compute anomalies relative to a running-mean climatology rather than a single static climatology is meant to account for any climate drift that may be present, under the assumption that the drift would be similar in the perturbed (abrupt4$\times CO_2$) and preindustrial control (piControl) runs. We have tested the impact of computing anomalies relative to the average over the entire piControl simulation and found averaging length to make little difference.

While linearity is a good first-order approximation, the slope of $\Delta N$ versus $\Delta T$ actually tends to decrease over time as the strength of individual feedback mechanisms are modulated by the background climate state (Armour et al. 2013; Rose et al. 2014). For some models this decrease is pronounced and induces error in our $F$ and $\lambda$ calculations (e.g., Andrews et al. 2012). We use the Gregory et al. (2004) approach despite this potential error because very few modeling groups provided sufficient data to compute feedbacks and forcings using other methods. We also note that the current study is focused entirely on intermodel spread, and Fig. 2a of Chung and Soden (2015b) indicates that spread in effective forcing is not strongly affected by the method used to derive it. Because our estimates of $F$ and $\lambda$ are calculated using the first 150 years of output after quadrupling $CO_2$, they probably do not represent true equilibrium values and are better thought of as estimates of the effective values over this period. In particular, the effective climate sensitivities calculated here are probably underestimates of the true equilibrium climate sensitivity (Williams et al. 2008; Winton et al. 2010; Andrews et al. 2015).

We use radiative kernels (Shell et al. 2008, hereafter Shell08; Soden et al. 2008, hereafter Soden08) to compute feedback terms for all components except Cld. In the kernel method, the radiative impact of a warming-induced change in factor $i$ is computed by multiplying the modeled change in $i$ by the TOA radiative response to a one-unit change in $i$:

$$\Delta R_i = \frac{\partial R}{\partial i}\Delta i. \tag{3}$$

This method is convenient because $\partial R/\partial i$ tends to be independent of model (Shell08) and of the magnitude of $i$ for reasonably small perturbations (Jonko et al. 2012; Block and Mauritsen 2013; Chung and Soden 2015a), so these quantities (called kernels) can be precomputed. For this study, we compute results using latitude-, longitude-, climatological monthly, and (where applicable) vertically resolved kernels from both Shell08 and Soden08. Monthly radiative anomalies and $\Delta i$ are computed for $i \in$ {Pl, WV, LR, Alb} by subtracting 21-yr running means of piControl run quantities from annually resolved samples taken from abrupt4$\times CO_2$ simulations. For feedbacks that depend on vertical structure (WV and LR), $\Delta R_i$ values are computed at each level, multiplied by layer mass, and summed over tropospheric levels (i.e., from the surface to the tropopause). We compute the tropopause for each time sample based on the World Meteorological Organization (1957) "threshold lapse rate" approach as implemented by Reichler et al. (2003).

Cloud feedback cannot be computed directly by multiplying the Soden08 or Shell08 kernels with anomalies in the model-diagnosed vertical profile of cloud fraction because the radiative impact of clouds at a given level is strongly affected by vertical overlap with clouds at other levels. Moreover, the TOA radiative impact of a change in cloud fraction depends on the cloud-top pressure and optical depth, neither of which is available as standard model output. There are two options to get around this problem: One is to use ISCCP simulator histograms of cloud fraction segregated by cloud-top pressure and cloud optical depth (Klein and Jakob 1999; Webb et al. 2001), along with similarly defined cloud radiative kernels (Zelinka et al. 2012). This allows for a direct calculation of the cloud feedback in a manner analogous to the kernel-derived calculations of the noncloud feedbacks. We do not use cloud kernels for this study because doing so requires ISCCP simulator data, which is not available for most of the models from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012). The other option, which we do implement, is to adjust the change in cloud radiative effect (CRE) to account for noncloud influences (Soden08; Shell08). CRE is defined as the difference between all-sky and clear-sky radiative fluxes at the TOA, so anomalies in CRE can be caused not only by cloud changes but also by changes in Pl, WV, LR, and Alb (Soden et al. 2004). Failure to remove the component of $\Delta$CRE related to other feedbacks results in double counting, causing net $\lambda$ to no longer be equal to the sum of its component feedbacks. To avoid this, we use the adjusted $\Delta$CRE, defined as

$$\Delta R_{\mathrm{Cld}} = \Delta \mathrm{CRE} - \sum_{i \neq \mathrm{Cld}} (\Delta R_i - \Delta R_i^0), \tag{4}$$

where superscript 0 indicates calculations using clear-sky variables. Adjusted $\Delta$CRE represents the radiative anomaly due solely to clouds. SW and LW cloud components are computed by applying Eq. (4) to SW or LW radiative components individually.

TABLE 1. Feedbacks and forcing used in this study. Models with names in boldface pass a clear-sky linearity test. Values in parentheses are computed relative to constant RH as advocated by HS12. All values are computed using the Soden08 kernels. (Expansions of acronyms are available at http://www.ametsoc.org/PubsAcronymList.)

| Model | $\lambda_{Pl}$ | $\lambda_{WV}$ | $\lambda_{LR}$ | $\lambda_{SW}$ | $\lambda_{LW}$ | $\lambda_{Alb}$ | $\lambda_{Re}$ | $\lambda$ | $F$ |
|---|---|---|---|---|---|---|---|---|---|
| ACCESS1.0 | −3.12 (−1.85) | 1.65 (0.04) | −0.49 (−0.13) | 0.11 | 0.39 | 0.39 | 0.30 | −0.77 | 2.95 |
| **ACCESS1.3** | −3.16 (−1.85) | 1.65 (0.01) | −0.49 (−0.14) | 0.53 | 0.15 | 0.37 | 0.13 | −0.82 | 2.90 |
| **BNU-ESM** | −3.08 (−1.86) | 1.46 (0.04) | −0.17 (0.03) | −0.12 | 0.36 | 0.56 | 0.07 | −0.92 | 3.72 |
| **CCSM4** | −3.14 (−1.88) | 1.47 (−0.02) | −0.29 (−0.03) | −0.00 | 0.27 | 0.43 | 0.06 | −1.18 | 3.48 |
| CNRM-CM5 | −3.13 (−1.85) | 1.55 (−0.02) | −0.38 (−0.09) | −0.16 | 0.36 | 0.41 | 0.21 | −1.14 | 3.70 |
| CSIRO Mk3.6.0 | −3.21 (−1.85) | 1.71 (−0.01) | −0.56 (−0.17) | 0.61 | 0.10 | 0.35 | 0.37 | −0.64 | 2.60 |
| **CanESM2** | −3.22 (−1.89) | 1.77 (0.08) | −0.49 (−0.13) | −0.24 | 0.78 | 0.36 | 0.00 | −1.03 | 3.82 |
| **FGOALS-g2** | −3.15 (−1.91) | 1.52 (0.08) | −0.22 (0.01) | 0.00 | 0.39 | 0.54 | 0.07 | −0.85 | 2.85 |
| FGOALS-s2 | −3.13 (−1.84) | 1.71 (0.09) | −0.45 (−0.12) | −0.33 | 0.44 | 0.49 | 0.37 | −0.91 | 3.81 |
| **GFDL CM3** | −3.12 (−1.82) | 1.67 (−0.06) | −0.64 (−0.21) | 0.64 | 0.28 | 0.40 | 0.04 | −0.74 | 2.96 |
| GFDL-ESM2G | −3.15 (−1.82) | 1.88 (0.06) | −0.81 (−0.29) | −0.34 | 0.62 | 0.26 | 0.32 | −1.22 | 2.97 |
| GFDL-ESM2M | −3.14 (−1.81) | 1.90 (0.03) | −0.88 (−0.32) | −0.47 | 0.62 | 0.24 | 0.36 | −1.37 | 3.35 |
| GISS-E2-H | −3.11 (−1.85) | 1.69 (0.06) | −0.53 (−0.13) | −0.71 | 0.60 | 0.27 | 0.11 | −1.66 | 3.82 |
| GISS-E2-R | −3.11 (−1.84) | 1.81 (0.12) | −0.62 (−0.18) | −0.78 | 0.65 | 0.22 | 0.07 | −1.77 | 3.74 |
| **HadGEM2-ES** | −3.13 (−1.84) | 1.62 (0.04) | −0.41 (−0.10) | 0.31 | 0.42 | 0.35 | 0.20 | −0.64 | 2.93 |
| **IPSL-CM5A-LR** | −3.19 (−1.83) | 2.02 (0.06) | −0.94 (−0.32) | 0.63 | 0.54 | 0.23 | −0.04 | −0.75 | 3.10 |
| **IPSL-CM5A-MR** | −3.22 (−1.83) | 2.10 (0.06) | −1.05 (−0.38) | 0.62 | 0.62 | 0.16 | −0.04 | −0.80 | 3.31 |
| **IPSL-CM5B-LR** | −3.12 (−1.86) | 1.67 (0.07) | −0.45 (−0.09) | 0.38 | 0.24 | 0.24 | 0.03 | −1.02 | 2.65 |
| **MIROC-ESM** | −3.26 (−1.87) | 1.94 (0.05) | −0.79 (−0.27) | 0.23 | 0.45 | 0.47 | 0.02 | −0.92 | 4.29 |
| **MIROC5** | −3.19 (−1.85) | 1.86 (0.08) | −0.67 (−0.21) | −0.32 | 0.32 | 0.42 | 0.04 | −1.53 | 4.15 |
| **MPI-ESM-LR** | −3.10 (−1.79) | 1.80 (−0.05) | −0.80 (−0.26) | −0.14 | 0.60 | 0.39 | 0.11 | −1.13 | 4.10 |
| **MPI-ESM-MR** | −3.11 (−1.79) | 1.83 (−0.06) | −0.87 (−0.30) | −0.07 | 0.56 | 0.37 | 0.08 | −1.19 | 4.11 |
| **MPI-ESM-P** | −3.12 (−1.80) | 1.85 (−0.04) | −0.87 (−0.29) | −0.19 | 0.61 | 0.36 | 0.12 | −1.23 | 4.27 |
| **MRI-CGCM3** | −3.19 (−1.91) | 1.49 (−0.05) | −0.34 (−0.05) | 0.30 | −0.00 | 0.45 | 0.06 | −1.22 | 3.19 |
| **NorESM1-M** | −3.11 (−1.87) | 1.48 (0.00) | −0.23 (−0.00) | 0.07 | 0.28 | 0.39 | 0.04 | −1.09 | 3.07 |
| **BCC_CSM1.1** | −3.15 (−1.88) | 1.42 (−0.05) | −0.22 (−0.01) | −0.09 | 0.39 | 0.42 | 0.06 | −1.15 | 3.26 |
| **BCC_CSM1.1(m)** | −3.21 (−1.89) | 1.54 (−0.05) | −0.32 (−0.05) | 0.04 | 0.38 | 0.36 | 0.01 | −1.19 | 3.46 |
| **INM-CM4.0** | −3.11 (−1.82) | 1.59 (−0.01) | −0.43 (−0.12) | 0.01 | 0.22 | 0.38 | −0.09 | −1.43 | 2.97 |

For each mechanism $i$, annually averaged $\Delta R_i$ is regressed on corresponding global- and annual-average $\Delta T$ values (as done for net $\lambda$ and $F$) to get $\lambda_i$ (from the slope) and the tropospheric adjustment of $i$ to $CO_2$ forcing $F_i$ (from the $y$ intercept). These tropospheric adjustments act to modify the radiative forcing of $CO_2$.

Error in the kernel method can be quantified by comparing the change in clear-sky TOA radiation computed as $\sum_i \Delta R_i^0$ against values extracted directly from the model. This is known as the clear-sky linearity test and is described further in Shell08 and Soden08. As noted by Jonko et al. (2012) and Block and Mauritsen (2013), using kernels derived from current-climate conditions can introduce substantial error when applied to $4 \times CO_2$ simulations. This is the reason that 8 of the 28 CMIP5 models with sufficient data for our calculations have discrepancies between kernel-derived and model-output TOA clear-sky flux changes of greater than 15% in the global average. We exclude these models from further analysis.

Calculated values for all models are presented in Table 1 (model data available online at https://pcmdi9.llnl.gov/projects/cmip5/; data accessed September 2015). As recommended by Vial et al. (2013, hereafter Vial13), we include the residual Re between net $\lambda$ and the sum of its components as an explicit term in our table. Because Shell08 and Soden08 kernels both provide very similar numbers and more models pass the clear-sky linearity test for Soden08 kernels, we use Soden08 values exclusively for the rest of this paper.

*Feedbacks using constant RH*

Traditionally, $\lambda_{WV}$ is taken to be the radiative impact of water vapor specific humidity $q_v$ increase, $\lambda_{LR}$ is the radiative change associated with vertically nonuniform atmospheric temperature increase, and $\lambda_{Pl}$ is the radiative effect of surface warming. Held and Shell (2012, hereafter HS12) note, however, that $q_v$ typically changes in such a way that relative humidity (RH) remains roughly constant. As a result, water vapor increases as atmospheric temperature increases (leading to the correlation between WV and LR feedbacks noted above). HS12 suggest defining the water vapor feedback based on change in $q_v$ associated with changes in relative humidity. To have all $\lambda_i$ sum to $\lambda$, the radiative impact associated with $q_v$ change at constant RH must then be partitioned such that the component responsible for warming each atmospheric column by its surface $\Delta T$ is added to $\lambda_{\rm Pl}$ and the impact of $q_v$ change at constant RH associated with nonuniform atmospheric warming is added to $\lambda_{\rm LR}$. This repartitioning of feedbacks has the advantage (shown below) of reducing covariance between $\lambda_{\rm WV}$ and $\lambda_{\rm LR}$. As explained in HS12, this new set of feedback definitions also avoids issues with computing feedbacks relative to a physically unrealizable (supersaturated) base state. We consider the impact of this feedback repartitioning in section 6.

## 3. DB08 method

There are two key steps to the DB08 process for identifying the main sources of ECS spread. The first step is to rewrite Eq. (2) as a sum of terms where each term is uniquely associated with a particular feedback or forcing mechanism. The second step is to compute the intermodel standard deviation of each summand and to use the result as a proxy for the importance of that term to ECS spread. In this section we provide an alternative derivation for the DB08 method of writing ECS as a sum of terms that is simpler and highlights some features of their approach. We start with Eq. (2) and note that $\lambda = \lambda_{\rm Pl} + \sum_{i\neq {\rm Pl}}\lambda_i$:

$$\begin{aligned}\mathrm{ECS} &= \frac{-F}{\lambda}\frac{\lambda_{\rm Pl}}{\lambda_{\rm Pl}}\\ &= \frac{-F}{\lambda}\frac{\left(\lambda - \sum\limits_{i\neq {\rm Pl}}\lambda_i\right)}{\lambda_{\rm Pl}}\\ &= \frac{-F}{\lambda_{\rm Pl}} - \frac{\mathrm{ECS}}{\lambda_{\rm Pl}}\sum_{i\neq {\rm Pl}}\lambda_i. \end{aligned} \tag{5}$$

DB08 then breaks $F$ into multimodel mean $\overline{F}$ and perturbation $F'$ components and expands Eq. (5):

$$\begin{aligned}\mathrm{ECS} &= -(\overline{F}+F')\left(\frac{1}{\lambda_{\rm Pl}} - \sum_{i\neq {\rm Pl}}\frac{\lambda_i}{\lambda_{\rm Pl}}\frac{1}{\lambda}\right)\\ &= \frac{-\overline{F}}{\lambda_{\rm Pl}} + \sum_{i\neq {\rm Pl}}\frac{\lambda_i}{\lambda_{\rm Pl}}\frac{\overline{F}}{\lambda} - \frac{F'}{\lambda_{\rm Pl}}\left(1-\frac{\lambda-\lambda_{\rm Pl}}{\lambda}\right)\\ &= \frac{-\overline{F}}{\lambda_{\rm Pl}} + \sum_{i\neq {\rm Pl}}\frac{\lambda_i}{\lambda_{\rm Pl}}\frac{\overline{F}}{\lambda} - \frac{F'}{\lambda}. \end{aligned} \tag{6}$$

DB08 partition ECS into $T$ contributions from individual feedback and forcing mechanisms by using the following definitions: $T_{\rm Pl} = -\overline{F}/\lambda_{\rm Pl}$, $T_i = \lambda_i\overline{F}/(\lambda_{\rm Pl}\lambda)$ for $i \notin \{{\rm Pl}, F\}$, and $T_F = -F'/\lambda$. In essence, all forcing variability is put into the $T_F$ term, then $T_{\rm Pl}$ is calculated using $\overline{F}$, and the remainder of ECS is partitioned among remaining feedbacks with size proportional to their relative magnitude.

In addition to being simpler than the derivation in DB08, our derivation clearly shows that the DB08 partitioning is mathematically nonunique. For example, $\lambda_{\rm Pl}/\lambda_{\rm Pl}$ on the first line in Eq. (5) could be replaced with $\lambda/\lambda$ or any other expression of the form $x/x$ to produce alternative ECS partitionings. Fundamentally, this nonuniqueness stems from the fact that moving the summation from the denominator to the numerator of Eq. (2) is ill posed. In general, any linear combination of terms that sum to ECS for each model is a mathematically acceptable partitioning. Some partitionings are more physically defensible than others, however. The DB08 partitioning is physically attractive because $T_{\rm Pl}$ is the temperature response needed to balance $\overline{F}$ in the absence of other feedbacks, and for $i \notin \{{\rm Pl}, F\}$, $T_i$ is the Planck response needed to balance the radiative perturbation caused by feedback $i$. But DB08 is not the only reasonable partitioning and it has some strange attributes. For example, since $T_i$ for $i \notin \{{\rm Pl}, F\}$ is constructed by multiplying $\lambda_i$ by $\overline{F}/(\lambda_{\rm Pl}\lambda)$ and this latter quantity itself has substantial intermodel spread, $\mathrm{var}(T_i)$ will be nonzero even if $\mathrm{var}(\lambda_i) = 0$. Further, any $\lambda_i \in \{{\rm Pl}, F\}$ with important intermodel variations will drive coherent variations in $\lambda$, which will damp the importance of $i$ since its corresponding $T_i$ is proportional to $\lambda_i/\lambda$.

## 4. A better partitioning

The above discussion leaves us with a few obvious questions: Is there a better way to decompose ECS spread? Do the details of partitioning strategy matter? In this section we describe a new partitioning strategy that avoids the issues noted above. By comparing this new approach against that from DB08, we are able to show in section 6 that partitioning strategy does have a large impact on our understanding of which terms are important for ECS spread.

Our new approach is based on linearizing $1/\lambda$ about its multimodel mean value using a Taylor expansion:

$$\begin{aligned}\mathrm{ECS} &= \frac{-F}{\overline{\lambda}+\lambda'}\\ &= \frac{-F}{\overline{\lambda}}\sum_{j=0}^{\infty}\left(\frac{-\lambda'}{\overline{\lambda}}\right)^j\\ &= \frac{-F}{\overline{\lambda}}\left[1-\frac{\lambda'}{\overline{\lambda}}+\left(\frac{\lambda'}{\overline{\lambda}}\right)^2\sum_{j=0}^{\infty}\left(\frac{-\lambda'}{\overline{\lambda}}\right)^j\right]\\ &= \frac{-\overline{F}-F'}{\overline{\lambda}}\left(1-\frac{\lambda'}{\overline{\lambda}}+\frac{\lambda'^2}{\overline{\lambda}\lambda}\right)\\ &= \frac{-\overline{F}}{\overline{\lambda}} - \frac{F'}{\overline{\lambda}} + \frac{\overline{F}}{\overline{\lambda}^2}\sum_i\lambda'_i + \frac{F'\lambda'}{\overline{\lambda}^2} - \frac{F\lambda'^2}{\overline{\lambda}^2\lambda}. \end{aligned} \tag{7}$$

The second line in Eq. (7) results from replacing $1/(1+\lambda'/\overline{\lambda})$ with a Taylor expansion; the fourth line results from applying the same identity in reverse.

Nonlinearity in the system is contained entirely in the last two terms of Eq. (7); we combine these terms into a single quantity that we call $T_{\mathrm{nonlin}}$. We show in section 6 that $\mathrm{var}(T_{\mathrm{nonlin}})$ is negligibly small.

The first term in Eq. (7) is large and is not related to any particular feedback mechanism. This makes linearization inappropriate for determining which processes are important to the multimodel mean magnitude of ECS. Because our paper focuses entirely on ECS spread, this term is not important since it is constant across models and therefore makes no contribution to var(ECS).

The second term in Eq. (7) is the contribution to ECS spread from spread in effective forcings $T_F$. The third term in Eq. (7) is the ECS contribution from $\lambda_i$: $T_i = \overline{F}/\overline{\lambda}^2\lambda_i'$. Because the only quantity in this last term with intermodel spread is $\lambda_i'$, our decomposition avoids the problem of having variance that is assigned to one feedback but is influenced by another.

Since we have linearized the contribution of each mechanism to var(ECS), the new partitioning is only an approximate partitioning of ECS. It is therefore incumbent on us to prove that linearization does not significantly distort our understanding of the main processes contributing to ECS spread. This issue is addressed in detail in the appendix. To summarize, the $\lambda_i$ with largest variance is not necessarily the largest contributor to spread in $1/\sum_i\lambda_i$, but linearization is adequate for $\lambda'_i \ll \overline{\lambda_i}$. Only $\lambda_{\mathrm{SW\_Cld}}$ perturbations are large enough for linearization to induce substantial error in our dataset (and most likely in other CMIP-class ensembles). Since shortwave cloud feedback is shown in section 6 to be the largest contributor to intermodel spread and the appendix shows that neglecting higher-order terms acts to damp the importance of terms with large perturbations, linearization does not affect our conclusions and is therefore appropriate here.

## 5. The variance of a sum

Once ECS has been written as a sum of terms, its variance can be decomposed into the variances and covariances of those terms by using the following identity:

$$\mathrm{var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{cov}(X_i, X_j)$$
$$= \sum_{i=1}^{N}\mathrm{var}(X_i) + 2\sum_{j=1}^{N}\sum_{k=j+1}^{N}\mathrm{cov}(X_i, X_j). \quad (8)$$

By focusing only on the intermodel standard deviation of summands, DB08 and Vial13 missed potentially important covariance terms.

Plotting the $T_i$ covariance matrix is a nice way to visualize the contributions of forcing and feedback processes to ECS spread. This is done in Fig. 1. Variances for each process appear on the main diagonal, while covariance terms are on the off diagonals. Because the covariance matrix is symmetric, each covariance should be included twice. For ease of display, we instead double the value of covariances above the main diagonal and omit the corresponding covariances below the diagonal. Covariance terms can be positive or negative while variances are always nonnegative. For ease of display, each matrix in Fig. 1 has been normalized by var(ECS) ($=0.475\,\mathrm{K}^2$) such that the sum of all its entries is 1.

## 6. Results

First, note that covariance and variance terms in Fig. 1 are of comparable magnitude regardless of partitioning approach; this implies that ignoring covariance terms is inappropriate.

Next, note that the DB08 and new partitioning methods yield extremely different covariance matrices (Figs. 1a and 1c, respectively). Variance and covariance tend to be much more homogeneously distributed in DB08 because $T_i \propto \lambda_i/\lambda$ for $i \notin \{\mathrm{Pl}, F\}$. Dividing by $\lambda$ damps contributions from the largest-variance contributors (which drive sympathetic variations in $\lambda$) and adds variance to terms that have little on their own. Using $\lambda_i$ rather than $\lambda'_i$ in combination with normalization by $\lambda$ also contributes significantly to differences between the DB08 and new partitionings (not shown).

It is also worth pointing out that the rows and columns in Fig. 1c related to feedbacks are identical to the covariance matrix for variance in net feedback (subject to rescaling) because $T_i \propto \lambda'_i$ in our new approach. In other words, our new method is equivalent to simply assuming that the feedbacks with largest variance are the biggest sources of ECS spread (as done by Bony et al. 2006; Soden and Held 2006; Huybers 2010). Our method extends these previous approaches by including forcing terms and by providing mathematical justification for this approximation.

One troubling aspect of Fig. 1a is that its variance terms seem different than presented by DB08 for CMIP3 models and by Vial13 for CMIP5 models. In particular, both DB08 and Vial13 found $\lambda_{\mathrm{Cld}}$ to be the dominant source of intermodel spread, yet $\lambda_{\mathrm{SW\_Cld}}$ and $\lambda_{\mathrm{LW\_Cld}}$ in the current study both have magnitude smaller than $F$ and $\lambda_{\mathrm{WV}}$. While some differences are expected as a result of using different models and computing variance instead of standard deviation, the main reason for inconsistency with previous studies is that DB08 and Vial13 presented only the combined

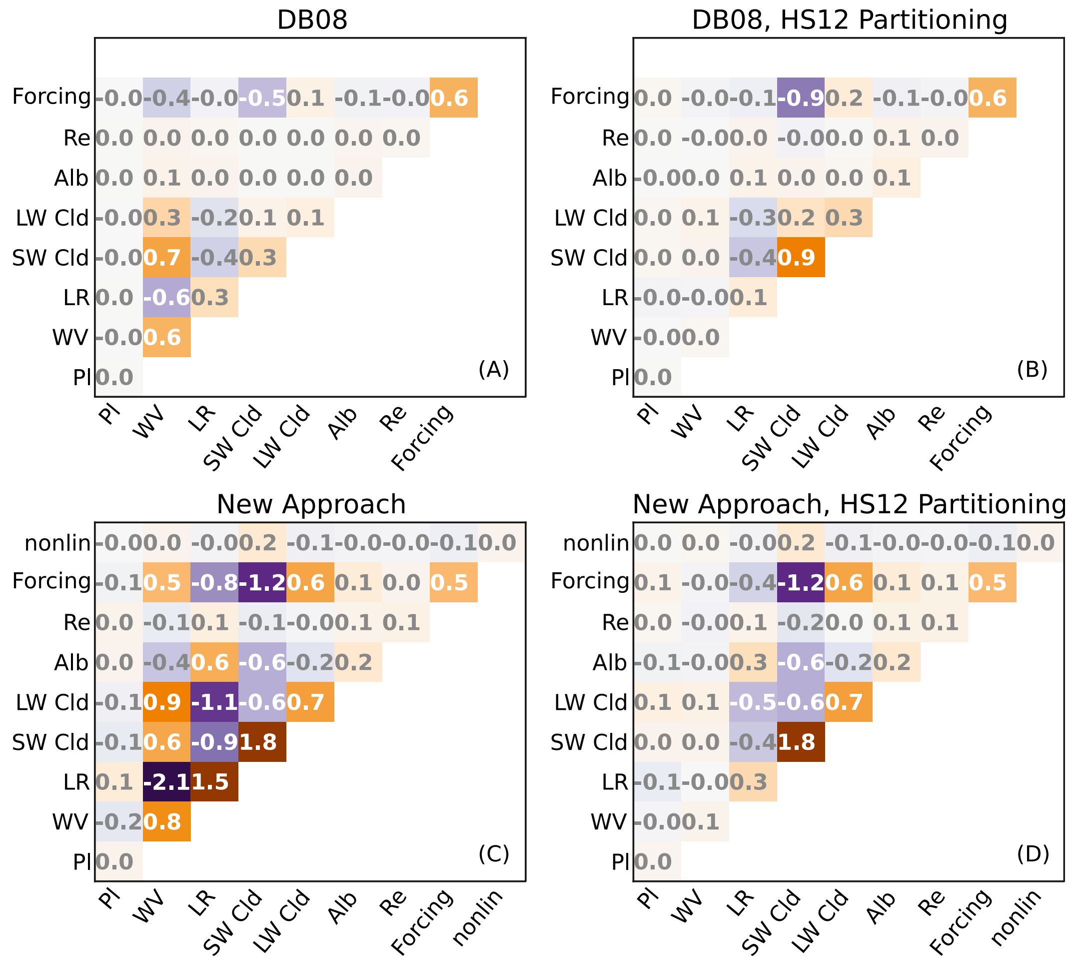


FIG. 1. Fractional contributions of partitioned terms to the total variance of ECS. Colors indicate magnitude, which is also given by numbers in each cell.

LW + SW cloud feedbacks and WV + LR feedbacks rather than individual components. Looking at Fig. 1a, we see that LW and SW Cld contribute 0.1 and 0.3 fractional units to the total spread (respectively) and their covariance terms add an additional 0.1 unit, so the variance contributed by net feedback sums to 0.5 units. Combining WV and LR in a similar way yields a spread of only 0.3 units because WV and LR variations compensate each other strongly. Thus after combining terms $\lambda_{\mathrm{Cld}}$ again stands out as a dominant source of intermodel spread in ECS. This can also be seen in Fig. 2, which shows feedbacks after combining SW with LW Cld and WV with LR. So is it better to consider these terms separately or in combination? Combining terms simplifies interpretation but reduces information content, so the answer to this question probably depends on the application of interest.

Even after accounting for combined terms, Fig. 1a indicates greater importance for var($F$) than suggested by Vial13. This can be understood by examining the forcing terms in Eqs. (25) and (26) of Vial13. These equations are analogous to our Eq. (5), which is an intermediate stage where forcing terms have not been isolated yet. Isolating forcing variability into a single term [as done in our Eq. (6) and DB08's Eq. (15)] results in the forcing term being normalized by $\lambda$ rather than $\lambda_{\mathrm{Pl}}$. Since $\lambda_{\mathrm{Pl}} \approx 3\lambda$, this difference greatly (and inappropriately) reduces their estimates of intermodel variance in $F$.

Figures 1b,d show the partitioning of intermodel spread using the HS12 definitions for WV, LR, and Pl. As expected, the HS12 approach reduces compensation between WV and LR and reduces intermodel variance in these terms as a result. Even with HS12 definitions, however, covariance terms remain nonnegligible. Switching to HS12 definitions has a large effect on all terms when the DB08 partitioning is used because $\lambda_{\mathrm{Pl}}$ is used as a scaling factor for all $\Delta T_i$. In the new ECS

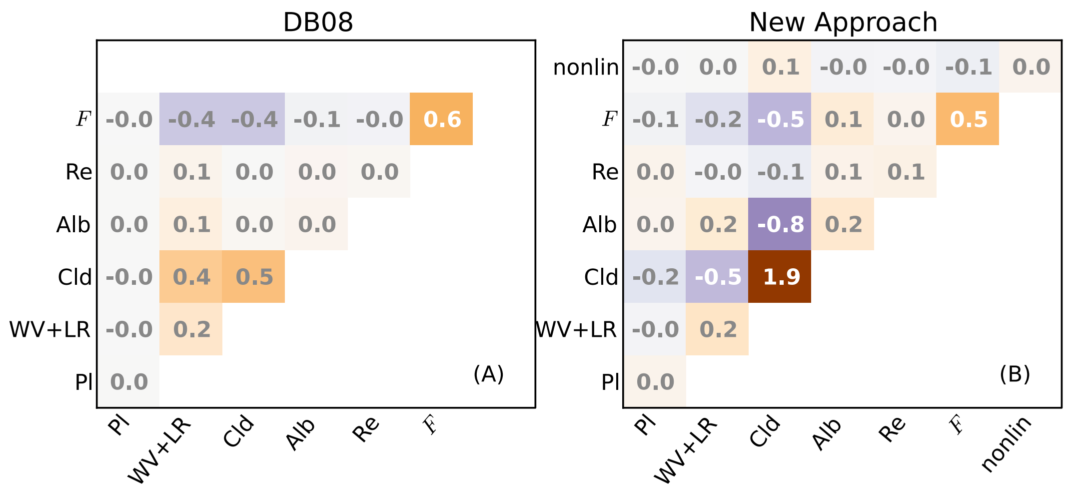


FIG. 2. As in Fig. 1, but using standard feedback definitions only and combining $\lambda_{\mathrm{SW\_Cld}}$ with $\lambda_{\mathrm{LW\_Cld}}$ and $\lambda_{\mathrm{LR}}$ with $\lambda_{\mathrm{WV}}$.

partitioning, feedbacks other than Pl, WV, and LR are unchanged (as expected). Sources of intermodel spread are more consistent between DB08 and our new method when using these new feedback definitions. Reduced ambiguity regarding the source of ECS variability is another good reason to use the HS12 definitions.

Because HS12 definitions are based in sound physics and reduce discrepancies between partitioning methods, we adopt this approach for our analysis of the sources of ECS spread. Note that the nonlinear term in Fig. 1d is negligibly small, a necessary condition for our new approach to be effective. Having a small nonlinear term related to the net feedback is not sufficient proof that individual feedbacks can be linearized, but our analysis in the appendix shows this assumption to be appropriate for our dataset. The residual term $T_{\mathrm{Re}}$, which measures error in the radiative kernels, is also acceptably small.

By far the largest term in Fig. 1d is var($\lambda_{\mathrm{SW\_Cld}}$). This is consistent with previous papers (e.g., Cess et al. 1990, 1997; Webb et al. 2013), which all show that SW cloud feedback is the dominant source of uncertainty in climate sensitivity. The second-largest term in Fig. 1d is covariance between $\lambda_{\mathrm{SW\_Cld}}$ and $F$; Ringer et al. (2014) also noted that anticorrelation between forcing and cloud feedback acts to strongly damp intermodel spread in ECS. Other important terms are var($\lambda_{\mathrm{LW\_Cld}}$), var($F$), cov($\lambda_{\mathrm{SW\_Cld}}$, $\lambda_{\mathrm{Alb}}$), cov($\lambda_{\mathrm{LW\_Cld}}$, $F$), cov($\lambda_{\mathrm{LW\_Cld}}$, $\lambda_{\mathrm{Alb}}$), and cov($\lambda_{\mathrm{SW\_Cld}}$, $\lambda_{\mathrm{Alb}}$). Huybers (2010) and Mauritsen et al. (2013) have previously noted compensation between $\lambda_{\mathrm{SW\_Cld}}$ and traditionally defined $\lambda_{\mathrm{WV+LR}}$. Our use of RH-based feedback definitions allows us to clarify that it is the lapse rate rather than moisture changes that are responsible for this relationship. Our separation of SW and LW Cld feedbacks makes clear that it is primarily LW rather than SW cloud feedbacks that are responsible for this behavior. Huybers (2010) and Mauritsen et al. (2013) also note negative cov($\lambda_{\mathrm{SW\_Cld}}$, $\lambda_{\mathrm{Alb}}$).

It is worth asking whether these covariances are robust or spurious. The best way to answer this question is to come up with an unassailable physical explanation for the relationship. A plausible explanation for the fairly strong negative correlation between LR and LW cloud feedbacks is that models with larger upper-tropospheric warming (and thus a more negative LR feedback) experience a more pronounced deepening of the troposphere (O'Gorman and Singh 2013), which allows high clouds to rise higher, thereby producing a larger LW cloud feedback (Zelinka and Hartmann 2010). Physical explanations for other covariances are still needed.

In the absence of a physical explanation, it is worth noting that $\mathrm{cov}(x, y) = [\mathrm{var}(x)]^{1/2}[\mathrm{var}(y)]^{1/2}\mathrm{corr}(x, y)$. This means that a large covariance can result either from a robust underlying relationship (i.e., a large-magnitude correlation) or from a large variance amplifying a very weak correlation. This is explored in Fig. 3 by scatterplotting pairs of quantities with substantial covariance. In Fig. 3 we see that there is very little correlation between $T_{\mathrm{SW\_Cld}}$ and $T_{\mathrm{LW\_Cld}}$ and between $T_{\mathrm{SW\_Cld}}$ and $T_{\mathrm{LR}}$, even though these terms had sizable covariance. Relationships with large covariance but small correlation indicate sources of ECS variance due to noise in the ensemble; these relations will probably change from ensemble to ensemble and are not worth exploring. The opposite may also occur; highly correlated quantities may explain a very small fraction of var(ECS). This represents relationships that are real but not very important to climate sensitivity (and therefore of less interest). Such relationships receive small weighting in our covariance matrices. The other correlations in Fig. 3 are of moderate magnitude and warrant further exploration.

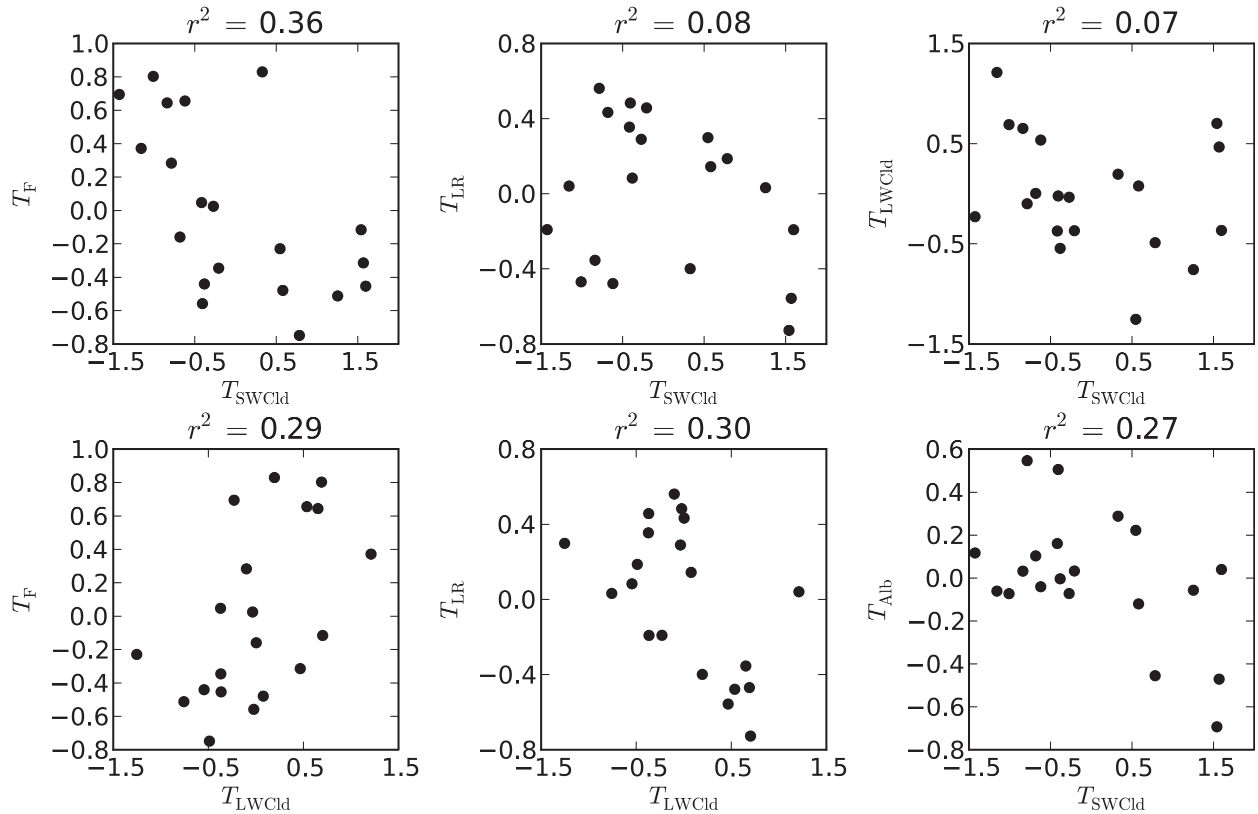


FIG. 3. Scatterplots of contributions to ECS shown in Fig. 1d to have substantial covariance. All plots are based on HS12 feedback definitions.

Another way to test whether results are robust is to repeat the analysis using different datasets. We have done this in a limited way by performing our analysis using a 10% cutoff for clear-sky kernel error (which yields 11 models instead of the 20 we use elsewhere) and using all available models regardless of clear-sky kernel error. Results look very similar in both cases (not shown), except cov($\lambda_{SW\_Cld}$, $\lambda_{Alb}$) becomes negligible when all models are included and cov($\lambda_{LW\_Cld}$, $\lambda_{LR}$) becomes small when a 10% clear-sky linearity test is adopted. This suggests that these covariances may be less robust, but further testing is needed to be definitive.

## 7. Conclusions

In this study we derive and analyze effective forcing and feedback information for a larger set of CMIP5 models than provided by earlier studies (Vial13; Andrews et al. 2012; Webb et al. 2013). We also explore an alternative way of defining water vapor feedback proposed by Held and Shell (2012). By recasting $\lambda_{WV}$ as the TOA radiative response to changes in relative rather than specific humidity, WV and LR feedbacks are decoupled. We show that this redefinition makes analyzing feedback mechanisms easier, so we hope that the HS12 technique is more widely adopted in the future.

Our main goal is to revisit the question of what causes disagreement in ECS between climate models. We note that the variance of a sum can be written as the sum of all possible combinations of covariances of summands. This insight leads us to show that covariance terms are as important to ECS spread as the more commonly analyzed variance terms. We also note that writing ECS as a sum of terms involving feedback processes is mathematically problematic because there is no good way to move $\sum_i \lambda_i$ from the denominator to the numerator of ECS $= -F/\sum_i \lambda_i$. As a result, there is no clear best way to decompose ECS. We then show that the details of how ECS is decomposed have a big impact on one's results. This leads us to think carefully about how to break ECS into a sum of terms. We point out that the DB08 approach does not fully isolate the contribution from a given feedback in the ECS contribution ascribed to that feedback. We propose a new method (based on linearizing $1/\lambda$) that does. We show

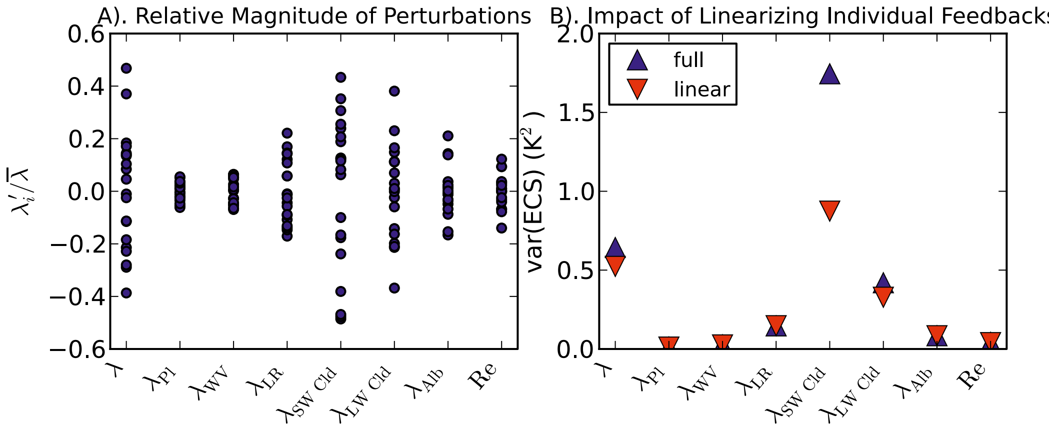


FIG. A1. (a) Perturbations normalized by $\overline{\lambda}$ for the terms listed on the $x$ axis; (b) var(ECS) computed after replacing all terms except the one listed on the $x$ axis with multimodel means. Blue triangles use the original full-complexity equation while red triangles use linearization. All panels use HS12 definitions.

in the appendix that linearization does not unduly distort the contributions from each feedback process to ECS spread.

Our results confirm previous work showing that $\lambda_{\mathrm{SW\_Cld}}$ is by far the dominant source of feedback spread. We also find that var($\lambda_{\mathrm{LW\_Cld}}$), var($F$), cov($\lambda_{\mathrm{SW\_Cld}}$, $F$), cov($\lambda_{\mathrm{LW\_Cld}}$, $F$), cov($\lambda_{\mathrm{LW\_Cld}}$, $\lambda_{\mathrm{LR}}$), and cov($\lambda_{\mathrm{SW\_Cld}}$, $\lambda_{\mathrm{Alb}}$) also contribute substantially to ECS spread. Understanding the physical mechanisms behind these sources of spread and improving their representation in climate models is the most direct path toward reducing uncertainty regarding the magnitude of global warming for a given increase in $CO_2$. The main contribution of this article is to point out that interactions between processes are themselves responsible for a great deal of ECS uncertainty and should be studied more closely. Since this point has not been widely appreciated in the past, there is probably great opportunity for advances in this area.

*Acknowledgments.* We would like to acknowledge the modeling groups [the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the World Climate Research Programme's Working Group on Coupled Modelling] for their roles in making available the CMIP5 multimodel dataset. Thanks also go to Thomas Reichler for making his WMO tropopause code publicly available. Support for these datasets is provided by the U.S. Department of Energy (DOE) Office of Science. This work was supported by the Office of Science (BER) at Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. All authors were supported by BER's Regional and Global Climate Modeling (RGCM) Program.

## APPENDIX

### Appropriateness of Linearizing 1/λ

If $|\lambda'| < |\overline{\lambda}|$, $1/\lambda$ can be expanded in a geometric series (or equivalently, a Taylor series):

$$\frac{1}{\lambda} = \frac{1}{\overline{\lambda}} \sum_{k=0}^{\infty} \left( \frac{-\lambda'}{\overline{\lambda}} \right)^k . \tag{A1}$$

Since the $(k + 1)$th summand is $|\lambda'/\overline{\lambda}|$ times as large as the $k$th summand, nonlinear terms can be neglected if $|\lambda'/\overline{\lambda}| \ll 1$. This is investigated in the first column of Fig. A1a, which shows $\lambda'/\overline{\lambda}$ for each model. This quantity is less than 0.2 for most models (indicating that linearization is appropriate), but seven models have relative perturbations greater than 0.2 and three have perturbations greater than 0.3. Because linearization error is proportional to the size of $\lambda$ perturbations, ensembles with greater diversity of $\lambda$ will experience more distortion by our method. Given the large size of linearization errors for some models, it is worth asking whether a linearization-based method is appropriate. We argue that our method is still worthwhile because it avoids the pathological issues of DB08 and the ambiguous interpretation of Webb et al. (2013), it induces small errors for most models, and it is the only existing method whose errors can be rigorously quantified. Additionally, the first column of Fig. A1b shows that linearization errors do not project strongly onto var(ECS). As a result, our decomposition works well for the data we have despite its linearity assumption. Finally, we show below that linearization error only affects the representation of processes with large perturbations, and its effect is to damp the importance of these dominant terms without changing the ordering of which terms are most important.

Since our goal is to obtain a qualitative rather than quantitative understanding of which terms are most important, we believe that our approach is useful despite its deficiencies.

Linearity of the net feedback does not ensure that the relative importance of individual feedbacks can be assessed by linearizing $1/\lambda$. In particular, a key feature of the linearized approach is that feedbacks with larger variance contribute more strongly to var(ECS). We will show via counterexample that such behavior is not always correct. Let $z = x + y$, where $x' = [-11.1, 11.1, -11.1, 11.1]$, $y' = [10.1, -10.1, 10.1, -10.1]$, and $\overline{z} = 10$. Anticorrelation between $x$ and $y$ has allowed us to make $[\mathrm{var}(x)]^{1/2}$ and $[\mathrm{var}(y)]^{1/2}$ comparable in magnitude to $\overline{z}$, even though $z'/\overline{z}$ is small enough for $1/z$ to be linearizable. For this example we will use the following equations as proxies for the importance of $x$ and $y$, respectively, to $\mathrm{var}(1/z)$:

$$\mathrm{var}\left(\frac{1}{x+\overline{y}}\right) = \mathrm{var}\left(\frac{1}{x'+\overline{x}+\overline{y}}\right) = \mathrm{var}\left(\frac{1}{x'+\overline{z}}\right) \tag{A2}$$

and

$$\mathrm{var}\left(\frac{1}{\overline{x}+y}\right) = \mathrm{var}\left(\frac{1}{\overline{x}+\overline{y}+y'}\right) = \mathrm{var}\left(\frac{1}{y'+\overline{z}}\right). \tag{A3}$$

Since $\mathrm{var}[1/(x'+\overline{z})] = 0.27$ and $\mathrm{var}[1/(y'+\overline{z})] = 25.25$, we see that $y$ is more important to $\mathrm{var}(1/z)$ than $x$, even though $\mathrm{var}(x) = 11.1^2 > \mathrm{var}(y) = 10.1^2$. This nonlinear behavior occurs because some values of $y' + \overline{z}$ are very close to zero, resulting in very large values of $1/(y' + \overline{z})$.

Mathematically, the reason why our counterexample behaved contrary to our linearity-based expectations is that $x'$ and $y'$ are of similar magnitude to $\overline{z}$, so $1/(x' + \overline{z})$ and $1/(y' + \overline{z})$ cannot be linearized. This prompts us to use the assumption that $\lambda'_i/\overline{\lambda} \ll 1$ as an additional requirement for using linearization to identify sources of var(ECS). Values for $\lambda'_i/\overline{\lambda}$ are provided in Fig. A1a. Relative perturbations for noncloud terms are all smaller than ~0.2, so linearizing these terms is appropriate. Cloud terms can be up to 0.5 units in magnitude, suggesting that linearization may induce significant distortion. We investigate the magnitude of this distortion in Fig. A1b by plotting var(ECS) after replacing all but one term with its multimodel mean value using the full nonlinear calculation and using the linearized calculation. Linearization has minimal impact on variance for all terms except $\lambda_{SW\_Cld}$, where linearization causes underprediction of ECS variance by a factor of 2. It makes sense that $\lambda_{SW\_Cld}$ has the largest error since it has the largest perturbations. It also makes sense that neglecting nonlinearity results in an underprediction of ECS variance since positive $\lambda_{SW\_Cld}$ perturbations bring $\overline{\lambda} + \lambda'_{SW\_Cld}$ very close to zero (resulting in large values of the reciprocal when nonlinearity is included). Since linearization damps the importance of terms proportionally to the size of that term's perturbations, it will damp the largest terms relative to the smaller terms without changing the ordering of each term's importance. In our case, only $\lambda_{SW\_Cld}$ seems to be significantly damped by linearization, and it still shows up as the dominant source of ECS spread. As a result, we consider linearization errors to be acceptable (but worth keeping in mind) for studies like this that seek to qualitatively assess the source of ECS spread in climate models. Because each ensemble is different, applicability of linearization should always be confirmed before applying our technique to a new dataset.

## REFERENCES

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophys. Res. Lett.*, **39**, L09712, doi:10.1029/2012GL051607.

——, ——, and ——, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, **28**, 1630–1648, doi:10.1175/JCLI-D-14-00545.1.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2013: Time-varying climate sensitivity from regional feedbacks. *J. Climate*, **26**, 4518–4534, doi:10.1175/JCLI-D-12-00544.1.

Block, K., and T. Mauritsen, 2013: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled $CO_2$. *J. Adv. Model. Earth Syst.*, **5**, 676–691, doi:10.1002/jame.20041.

Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate*, **19**, 3445–3482, doi:10.1175/JCLI3819.1.

Cess, R., 1975: Global climate change: An investigation of atmospheric feedback mechanisms. *Tellus*, **27A**, 193–198, doi:10.1111/j.2153-3490.1975.tb01672.x.

——, and Coauthors, 1990: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res.*, **95**, 16 601–16 615, doi:10.1029/JD095iD10p16601.

——, and Coauthors, 1997: Comparison of the seasonal change in cloud-radiative forcing from atmospheric general circulation models and satellite observations. *J. Geophys. Res.*, **102**, 16 593–16 603, doi:10.1029/97JD00927.

Chung, E.-S., and B. J. Soden, 2015a: An assessment of direct radiative forcing, radiative adjustments, and radiative feedbacks in coupled ocean–atmosphere models. *J. Climate*, **28**, 4152–4170, doi:10.1175/JCLI-D-14-00436.1.

——, and ——, 2015b: An assessment of methods for computing radiative forcing in climate models. *Environ. Res. Lett.*, **10**, 074004, doi:10.1088/1748-9326/10/7/074004.

Dufresne, J.-L., and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *J. Climate*, **21**, 5135–5144, doi:10.1175/2008JCLI2239.1.

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866. [Available online at https://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_Chapter09_FINAL.pdf.]

Gordon, N. D., and S. A. Klein, 2014: Low-cloud optical depth feedback in climate models. *J. Geophys. Res. Atmos.*, **119**, 6052–6065, doi:10.1002/2013JD021052.

Gregory, J., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, **31**, L03205, doi:10.1029/2003GL018747.

Held, I. M., and B. Soden, 2000: Water vapor feedback and global warming. *Annu. Rev. Energy Environ.*, **25**, 441–475, doi:10.1146/annurev.energy.25.1.441.

——, and K. M. Shell, 2012: Using relative humidity as a state variable in climate feedback analysis. *J. Climate*, **25**, 2578–2582, doi:10.1175/JCLI-D-11-00721.1.

Huybers, P., 2010: Compensation between model feedbacks and curtailment of climate sensitivity. *J. Climate*, **23**, 3009–3018, doi:10.1175/2010JCLI3380.1.

Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu, 2012: Climate feedbacks in CCSM3 under changing $CO_2$ forcing. Part I: Adapting the linear radiative kernel technique to feedback calculations for a broad range of forcings. *J. Climate*, **25**, 5260–5272, doi:10.1175/JCLI-D-11-00524.1.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.*, **127**, 2514–2531, doi:10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2.

Mauritsen, T., R. G. Graversen, D. Klocke, P. L. Langen, B. Stevens, and L. Tomassini, 2013: Climate feedback and synergy. *Climate Dyn.*, **41**, 2539–2554, doi:10.1007/s00382-013-1808-7.

O'Gorman, P. A., and M. S. Singh, 2013: Vertical structure of warming consistent with an upward shift in the middle and upper troposphere. *Geophys. Res. Lett.*, **40**, 1838–1842, doi:10.1002/grl.50328.

Reichler, T., M. Dameris, and R. Sausen, 2003: Determining the tropopause height from gridded data. *Geophys. Res. Lett.*, **30**, 2042, doi:10.1029/2003GL018240.

Ringer, M. A., T. Andrews, and M. J. Webb, 2014: Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophys. Res. Lett.*, **41**, 4035–4042, doi:10.1002/2014GL060347.

Rose, B. E. J., K. C. Armour, D. S. Battisti, N. Feldl, and D. D. B. Koll, 2014: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophys. Res. Lett.*, **41**, 1071–1078, doi:10.1002/2013GL058955.

Shell, K. M., J. T. Kiehl, and C. A. Shields, 2008: Using the radiative kernel technique to calculate climate feedbacks in NCAR's community atmospheric model. *J. Climate*, **21**, 2269–2282, doi:10.1175/2007JCLI2044.1.

Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *J. Climate*, **19**, 3354–3360, doi:10.1175/JCLI3799.1.

——, A. J. Broccoli, and R. S. Hemler, 2004: On the use of cloud forcing to estimate cloud feedback. *J. Climate*, **17**, 3661–3665, doi:10.1175/1520-0442(2004)017<3661:OTUOCF>2.0.CO;2.

——, I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying climate feedbacks using radiative kernels. *J. Climate*, **21**, 3504–3520, doi:10.1175/2007JCLI2110.1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Vial, J., J.-L. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dyn.*, **41**, 3339–3362, doi:10.1007/s00382-013-1725-9.

Webb, M., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905–922, doi:10.1007/s003820100157.

——, F. H. Lambert, and J. M. Gregory, 2013: Origins of differences in climate sensitivity, forcing, and feedback in climate models. *Climate Dyn.*, **40**, 677–707, doi:10.1007/s00382-012-1336-x.

Williams, K. D., W. J. Ingram, and J. M. Gregory, 2008: Time variation of effective climate sensitivity in GCMs. *J. Climate*, **21**, 5076–5090, doi:10.1175/2008JCLI2371.1.

Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate*, **23**, 2333–2344, doi:10.1175/2009JCLI3139.1.

World Meteorological Organization, 1957: Meteorology, a three-dimensional science: Second session of the Commission for Aerology. *WMO Bull.*, **4**, 134–138.

Zelinka, M. D., and D. L. Hartmann, 2010: Why is longwave cloud feedback positive? *J. Geophys. Res.*, **115**, D16117, doi:10.1029/2010JD013817.

——, S. A. Klein, and D. L. Hartmann, 2012: Computing and partitioning cloud feedbacks using cloud property histograms. Part I: Cloud radiative kernels. *J. Climate*, **25**, 3715–3735, doi:10.1175/JCLI-D-11-00248.1.

——, ——, K. Taylor, T. Andrews, M. Webb, J. Gregory, and P. Forster, 2013: Contributions of different cloud types to feedbacks and rapid adjustments in CMIP5. *J. Climate*, **26**, 5007–5027, doi:10.1175/JCLI-D-12-00555.1.