The DOI for this manuscript is doi: 10.1175/JCLI-D-12-00631.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Feldl, N., and G. Roe, 2013: The nonlinear and nonlocal nature of climate feedbacks. J. Climate. doi:10.1175/JCLI-D-12-00631.1, in press.

The nonlinear and nonlocal nature of climate feedbacks

Nicole Feldl *

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

Gerard H. Roe

Department of Earth and Space Sciences, University of Washington, Seattle, Washington

* *Corresponding author address:* Nicole Feldl, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640.
E-mail: feldl@uw.edu

ABSTRACT

The climate feedback framework partitions the radiative response to climate forcing into contributions from individual atmospheric processes. The goal of this study is to understand the closure of the energy budget in as much detail and precision as possible, within as clean an experimental set-up as possible. For an aquaplanet simulation under perpetual equinox conditions, we account for rapid tropospheric adjustments to CO_2 and diagnose radiative kernels for this precise model set-up. We characterize the meridional structure of feedbacks, heat transport, and nonlinearities in controlling the local climate response. Our results display a combination of positive subtropical feedbacks and polar amplified warming. These two factors imply a critical role for transport and nonlinear effects, with the latter acting to substantially reduce global climate sensitivity. At the hemispheric scale, a rich picture emerges: anomalous divergence of heat flux away from positive feedbacks in the subtropics; nonlinear interactions amongst and within clear-sky feedbacks, which reinforce the pattern of tropical cooling and high-latitude warming tendencies; and strong ice-line feedbacks that drive further amplification of polar warming. These results have implications for regional climate predictability, by providing an indication of how spatial patterns in feedbacks combine to affect both the local and nonlocal climate response, and how constraining uncertainty in those feedbacks may constrain the climate response.

1. Introduction

The power of the climate feedback framework lies in its ability to reveal the energy pathways by which the system adjusts to an imposed forcing, such as an increase in atmospheric CO_2 concentration. These internal adjustments may include changes in physical processes that control the distribution of clouds, water vapor, sea ice, and the vertical structure of temperature, which in turn act to amplify or dampen the surface temperature response to the forcing; these are the climate feedbacks. Further, the system may also adjust by redistributing energy between different latitudes, either by atmospheric or oceanic transport, or both. Understanding the relative importance and effectiveness of these different pathways is crucial for predicting the climate response to a perturbation.

We begin by reviewing the underpinnings of feedback analysis (e.g., Roe 2009). Climate feedbacks are closely related to the change in top-of-atmosphere (TOA) net radiative flux between two climate states, ΔR, which can be written as a Taylor series expansion in global-mean surface temperature change, $\Delta \overline{T}_s$:

$$\Delta R = A + B\Delta \overline{T}_s + \mathcal{O}\Delta \overline{T}_s^2. \tag{1}$$

The terms in Equation 1 can represent global averages, or be functions of latitude or grid cell, and in principle, one could alternatively choose to make the expansion in local temperature change (e.g. Crook et al. 2011; Armour et al. 2012). The sign convention is such that a positive radiative flux warms the system. The first term on the right-hand side, A, includes the external forcing itself, along with all changes in the energy balance that are independent of surface temperature change (i.e., semi-direct effects, see Section 2b). We refer to A as the climate forcing. The second term, $B\Delta \overline{T}_s$, reflects radiative flux changes that are linearly dependent on the system response $\Delta \overline{T}_s$; these are the classical feedback processes. Here, the sign of the feedback term is negative when the system is stable (i.e., a net negative feedback). The third component, $\mathcal{O}\Delta \overline{T}_s^2$, represents higher-order terms, which may reflect nonlinearities within individual processes or nonlinear interactions amongst different processes. In general, ΔR may be accommodated by either local heat storage or by a change in the divergence of atmospheric or oceanic heat transport ($\sim \Delta(\nabla \cdot F)$). We only present equilibrium calculations, so local heat storage can be neglected. Therefore, for the remainder of this study $\Delta(\nabla \cdot F)$ and ΔR are interchangeable. Further, in the global mean $\Delta(\nabla \cdot F) = 0$, and the feedback and nonlinear term must balance the forcing. Finally, Equation 1 is commonly written in a simplified form, with the nonlinear term $\mathcal{O}\Delta\overline{T}_s^2$ assumed minor and neglected (e.g. Senior and Mitchell 2000; Gregory et al. 2004; Soden and Held 2006), though we expressly evaluate it herein.

The goal of this study is to understand the closure of the TOA energy balance in as much detail and precision as possible. Doing so allows us to characterize the relative importance of the four terms in Equation 1—heat transport, forcing, feedbacks, and the nonlinearities—in controlling the local climate response. We carefully diagnose the climate forcing, taking into account the semi-direct (i.e., temperature independent) response of the atmosphere to CO_2 changes, and we derive the linear part of the response (i.e., the feedbacks) using radiative kernels explicitly calculated for our precise model set-up. In addition, we run our experiment in an idealized aquaplanet model with perpetual equinox conditions and a mixed-layer ocean, which minimizes complexities in the results.

Equation 1 can be rewritten using notation more common to the climate-feedbacks literature, and the nonlinear term can be expressed as a residual $\mathcal{R}$:

$$\underbrace{\mathcal{R}}_{\text{residual}} = \underbrace{\Delta R}_{\text{transport}} - \underbrace{\left[\left(\mathop{}_{x} \lambda_x\right)\Delta\overline{T}_s + \Delta\widetilde{R}_f\right]}_{\text{combined feedback and forcing}} = \mathcal{O}\Delta\overline{T}_s^2. \tag{2}$$

As discussed later, due to the methodology we use for determining the cloud feedback, the residual in this study applies only to the clear-sky physics (Section 2c and Appendix A). In our results we interpret this residual as the nonlinear term. Although we have tried to be diligent in rooting out common approximations that would contribute artificially to the residual, some may linger, and we return to this point in later sections.

Recall that ΔR in Equation 2 is the change in net TOA radiative flux and that it must be equal to the change in convergence of horizontal atmospheric heat flux, $\Delta(\nabla \cdot F)$; we refer to this term as "transport" for convenience. The forcing $\Delta\widetilde{R}_f$ is equivalent to A in Equation 1, where for clarity the tilde has been introduced to indicate the inclusion of semi-direct effects (discussed in Section 2b). We replace B with $\sum_x \lambda_x$. In much of the climate literature λ_x are known as feedback parameters, which we adopt here for consistency with earlier work, though we note this departs from conventional definitions (e.g. Bode 1945; Schlesinger 1985; Roe 2009). Physically the terms in the λ_x series simply reflect the linear decomposition of changes in the TOA energy budget (x represents water vapor, surface albedo, cloud, Planck and lapse rate feedbacks). Bony et al. (2006) provide a comprehensive review of the various climate feedbacks relevant on interannual to multidecadal timescales, and we will elaborate on the individual terms in following sections.

By construction, the feedback framework only provides an approximation to the actual TOA radiative flux changes, and hence to climate sensitivity. As mentioned above, a goal of this study is to understand the degree of approximation, and to the extent possible, assign physical meaning to the structure of the nonlinear term. How important are nonlinearities for the local energy balance, and do they provide insights into understanding ubiquitous features of climate change, such as polar amplification? While a handful of studies have quantified the linear approximation with respect to magnitude of forcing (Colman et al. 1997; Colman and McAvaney 2009; Jonko et al. 2012), we are unaware of any that emphasize the spatial pattern of *interactions* amongst clear-sky feedbacks. Further, how well must the forcing be represented to evaluate the energy balance? This question is partly motivated by recent work that has demonstrated a narrowing of the intermodel-spread in cloud feedback when rapid tropospheric adjustments are counted as part of the forcing (e.g., Andrews and Forster 2008).

We are motivated by a need to understand the implications of nonlinear and nonlocal feedbacks on regional climate predictability. Our first objective is a precise quantification of nonlinear interactions between feedbacks, using our idealized aquaplanet simulation. We also present an independent evaluation of the nonlinearity, in order to add physical meaning to our characterization. Our second objective is to understand the relative importance of contributions due to feedbacks, meridional heat transport, nonlinearities, and forcing to the spatial pattern of warming. In particular, this allows us to assess how local processes (i.e., feedbacks) affect nonlocal responses via transport. In essence, we have extended the feedback framework, conventionally applied to deconstructing global climate sensitivity, in order to evaluate the role of nonlinearities and dynamical effects on local temperature change.

2. Analysis

a. Aquaplanet model

We employ the Geophysical Fluid Dynamics Laboratory Atmospheric Model 2 (GFDL AM2) in its aquaplanet configuration. We specify perpetual equinox and daily-mean solar zenith angle. The ocean is represented as a 20-m mixed layer. Sea ice is treated as infinitesimally thin; the ocean albedo is increased to 0.5 where surface temperature drops below 263 K, but no ice thermodynamics are present in the experimental set-up. The critical temperature for sea-ice formation was chosen to reproduce a realistic ice-line latitude, when compared to the modern climate. A full description of the AM2 is provided by the GFDL Global Atmospheric Model Development Team (2004). This idealized configuration allows us to cleanly isolate the atmospheric response to CO_2 in the absence of coupled ocean physics, land-ocean contrast, land-surface processes, seasonal and diurnal cycles, and aerosol forcing. Our perturbation is achieved by an instantaneous doubling of CO_2, and then by integrating the model out to equilibrium.

Figure 1 shows climatological surface temperature and outgoing longwave radiation (OLR) for control and perturbation experiments, as well as the differences, for the last ten years of our 30-year integration. For this model set-up, doubling CO_2 results in a global-mean temperature increase of 4.69 K, a climate sensitivity that sits slightly above the upper end of the IPCC AR4 "likely" range (Solomon et al. 2007) and of AR5 models (Andrews et al. 2012). The shape of the temperature response as a function of latitude is characterized by strong polar amplification; warming peaks at 11.5 K in high northern latitudes, more than twice the global-mean. For comparison, Hwang et al. (2011) find that Arctic warming ranges from 2 to 3 times the global mean for CMIP3 simulations. Maxima in OLR occur over the dry subtropics, and the global-mean OLR for the control run is 235 W m^{-2}, which is about 10% larger than April climatology provided by NOAA-CIRES Climate Diagnostics Center[1]. In response to CO_2 doubling, there is a strong equatorial peak in ΔOLR associated with a 16% decrease in cloud fraction in the tropical upper-troposphere (see Fig. 4b). In nature, as in more complex models, the meridional structure of annual-mean OLR is blurred by seasonal variations in the position of the intertropical convergence zone (ITCZ), and by zonal asymmetries due to land-ocean contrast. The choice of perpetual equinox conditions, which produces a permanent equatorial ITCZ, leads to a focusing of many of the climate fields about the equator, which will also become apparent when we examine the patterns of water vapor and cloud feedbacks. This is a trade-off: we gain a clear picture of the feedback patterns and their dynamical causes in this idealized model, but must be more cautious about a direct application of the results to nature.

b. *Determination of radiative forcing*

Previous feedback studies have commonly assumed a spatially uniform radiative forcing based on estimates of the global mean (e.g. Soden et al. 2008). However the pattern of radiative forcing can be strikingly nonuniform, as we will show. Since our goal in this study is to close the energy balance as nearly as possible, an updated approach is desired that accounts for this spatial variability and is exact to our experimental set-up. Various

[1] Available online at http://www.cdc.noaa.gov.

definitions of radiative forcing are discussed in Hansen et al. (2005). We consider two methods: stratosphere-adjusted, in which the stratosphere is allowed to adjust radiatively to the presence of the forcing agent; and fixed-SST forcing, in which the troposphere is allowed to adjust as well. For a feedback analysis, the latter is to be strongly preferred since it accounts for all changes in forcing that are independent of surface temperature change. In other words, it is closest to the definition of A in Equation 1. We describe each forcing approach in more detail below.

The first method, stratosphere-adjusted radiative forcing, is calculated offline from the GFDL radiative transfer code, following definitions provided in the IPCC Third Assessment Report (Appendix 6.1 of Ramaswamy et al. 2001; Hansen et al. 2005). Under this classical "fixed dynamical heating" framework, the stratosphere is allowed to adjust to the forcing prior to calculating the TOA flux change. In other words, changes in the downward flux from the stratosphere, as a result of stratospheric temperature change, are assumed to be part of the forcing. The resulting quantity is sometimes called the "adjusted" radiative forcing, and it is relevant for CO_2 perturbation experiments because the adjustment of the stratosphere is argued to be fast compared to both the tropospheric response and the lifetime of the forcing agents (Hansen et al. 2005). Once the stratosphere has adjusted to its new radiative-dynamical equilibrium, the change in flux at the tropopause and at the TOA are identical. The solid gray line in Figure 2a shows the stratosphere-adjusted radiative forcing. It has a global mean value of 3.4 W m^{-2} and, notably, varies by about a factor of two as a function of latitude. The spatial pattern of the forcing is controlled by variations in surface temperature and high-level cloudiness (Shine and Forster 1999). Adding CO_2 beneath a region of extensive climatological cloud cover has less impact on TOA radiative fluxes. Highest values are thus found in the warm, cloud-free subtropics.

The second method, fixed-SST forcing, focuses on $\Delta\widetilde{R}_f$ as the climate forcing applied to the system *independent of and prior to a surface temperature response*. This definition is spurred by recent modeling results that have demonstrated semi-direct, tropospheric adjustments in response to CO_2 (in addition to the direct radiative effect of the greenhouse gas itself), which precede substantial surface warming and affect the TOA radiation balance. In particular, several studies (e.g., Andrews et al. 2011) have emphasized the importance of the cloud response operating over timescales less than one month. This rapid cloud adjustment manifests primarily as a shortwave effect of <1 W m^{-2}, which Colman and McAvaney (2011) suggest is driven by a decrease in relative humidity and cloud fraction in regions of enhanced heating at mid-to-lower levels in the troposphere. Other hypotheses involve shoaling of the planetary boundary layer due to suppressed surface heat fluxes (Watanabe et al. 2011) or reductions in entrainment (Wyant et al. 2012). Since it does not constitute a response to surface temperature change, any effect of rapid tropospheric adjustment is more properly combined with the forcing term. Failure to take this rapid adjustment into account as a forcing may bias the cloud feedback calculation.

We therefore perform a fixed-SST experiment, which is able to incorporate the rapid tropospheric adjustment to CO_2 prior to surface temperature change—in essence, turning off the feedbacks. A general critique of fixed-SST experiments in standard GCM configurations is that warming still occurs over land surfaces and sea ice, undermining the goal of having no surface response. However the aquaplanet integrations do not suffer from this inconsistency. We can easily fix surface temperature everywhere, and in effect equate the fixed-SST forcing of Hansen et al. (2005) with the "adjusted troposphere and stratosphere forcing" of Shine et al. (2003). The fixed-SST experiment is integrated for 40 years with zonally-symmetric and symmetric-about-the-equator specified SSTs (taken from the final year of our control run). It is otherwise identical to our model set-up for the feedback analysis. The forcing is then simply the change in net TOA radiative flux between $1{\times}CO_2$ and $2{\times}CO_2$ scenarios, with the first year discarded.

The solid black line in Figure 2a shows the climate forcing $\Delta\widetilde{R}_f$, including both external forcing and rapid tropospheric adjustments. It has a global-mean value[2] of 3.8 $\pm$

[2]The standard error (i.e., standard deviation of the mean) of the fixed-SST forcing is calculated from the monthly change in net TOA radiative flux after the doubling of CO_2. The estimated degree of uncertainty

$0.2\ \mathrm{W\ m^{-2}}$, close to that of the uniform forcing (Myhre et al. 1998). The fixed-SST and stratosphere-adjusted forcings share some similarities, particularly in the southern hemisphere, with maxima in the subtropics. However the fixed-SST forcing is characterized by notable, and perhaps surprising, hemispheric asymmetries. In fact, these asymmetries reflect exactly the physics we intended to capture in the forcing estimate. The clear-sky forcings (solid lines in Fig. 2b) are quite similar for both methods and hence it is the shortwave response of clouds to CO_2 (hashed line), which explains the variability, consistent with the proposed rapid cloud adjustment. In addition to the noisiness of the calculation, some of the hemispheric asymmetry may also be due to the perpetual equinox conditions that limit interaction between the hemispheres.

In the analysis that follows we predominantly use this fixed-SST forcing because it is nearest to our definition of a temperature-independent forcing, as presumed by the feedback framework, and because we believe it represents genuine variability in the forcing. The relatively small impact of this hemispheric asymmetry in forcing can be gauged from our results (see Fig. 3) and will be discussed in more detail in later sections; the differences also serve as a rough indication of how uncertainty in forcing influences the meridional structure of feedbacks.

c. *Kernels and feedbacks*

We apply the radiative kernel method of calculating climate feedbacks, following Soden and Held (2006) and Soden et al. (2008). The kernel represents the TOA radiative adjustment due to a differential nudge in the climate fields, and is calculated separately for changes in temperature, water vapor, and surface albedo. It can be thought of as a sensitivity matrix. A strength of our analysis is that we explicitly calculate radiative kernels for our precise experimental set-up, thus removing one of the most commonly-cited ambiguities associated

for the 40-year integration is comparable to values cited by previous studies (e.g., $0.3\ \mathrm{W\ m^{-2}}$ in Shine et al. 2003).

with this method (i.e., a mismatch between models used in kernel-generation and feedback-calculation, as occurs in intermodel comparisons; Zelinka and Hartmann 2012).

Radiative kernels are not the only approach for calculating feedbacks, and a comparison of various techniques can be found in Yoshimori et al. (2011). Briefly, kernels are a popular choice for intermodel comparisons because the calculation is based on a small and arguably non-model-specific perturbation (Soden and Held 2006), though they break down for sufficiently different mean states, such as under CO_2 octupling (Jonko et al. 2012). Non-kernel feedback calculations include partial radiative perturbation (PRP) and regression. The PRP method (Wetherald and Manabe 1988; Colman 2003) suffers from computational expense. The regression method of Gregory et al. (2004) is complicated by ambiguities associated with transient adjustments that can result in a poorly-constrained (or even misdiagnosed; Armour et al. 2012) feedback estimate, particularly when local scales are of interest, and by the inability to separately evaluate temperature, water vapor, and surface changes. Finally, recent studies have also proposed to reformulate the kernel framework around relative humidity, rather than specific humidity, thus removing the correlation between water vapor and lapse rate changes (Held and Shell 2012; Ingram 2012). However this rearrangement of energy flux changes into different individual feedbacks does not affect the total linear feedback nor the characterization of the nonlinear term, which is the focus of the present study.

Kernels show particular promise where nonlinear interactions are of interest. All feedback methods seek to characterize the linear decomposition of TOA radiative flux changes into the relative contributions from different physical processes. The PRP method is arguably the most exact decomposition of the differences between two climate states because the total (i.e., discrete) changes are used in the radiative calculations. However given our goal to estimate the linearity of climate feedbacks, the kernel method, in its use of small differential changes, is actually closer to the "tangent linear" approximation that is the formal basis for the Taylor series expansion in Equation 1.

Hence, following Soden and Held (2006) and Soden et al. (2008), we compute all feedbacks (with the exception of clouds) as products of two factors. The first is the change in TOA flux due to a small perturbation in variable x, and the second is the change in x between the two equilibrium climate states ($1\times CO_2$ and $2\times CO_2$), divided by the global-mean surface temperature response.

$$\lambda_x = \frac{\partial R}{\partial x} \cdot \frac{dx}{d\overline{T}_s} \tag{3}$$

where $\partial R/\partial x = K_x$ (i.e. the radiative kernel for x) and x represents temperature, specific humidity, and surface albedo. To create the kernels, instantaneous temperatures T, including the surface temperature T_s, are perturbed by 1 K; surface albedo α is perturbed by 1%; and specific humidity q is perturbed to match the change in saturation specific humidity that would occur from a 1 K warming, assuming fixed relative humidity. We perturb T, α, and q from the control climate for each latitude, longitude, time, and pressure level. The kernels are calculated from one year of instantaneous eight-times daily model output, using the offline radiation code. We make computations for clear skies (i.e., clouds instantaneously set to zero) as well as for all-sky conditions simulated by the model. The kernels we derive broadly resemble the kernels calculated from more realistic climate models (i.e., with land, seasonal cycles, etc.), as presented for instance in Soden et al. (2008). However the simplicity of our aquaplanet set-up means the spatial patterns of the kernels are sharper, and can be very clearly related to individual aspects of the atmospheric response. The kernels are presented and described in detail in Appendix B.

Feedbacks are calculated by convolving 10 years of equilibrated monthly anomalies with the 12-month kernels, in the case of temperature, water vapor, and albedo (Eqn. 3). The two parts comprising the temperature feedback are calculated from the surface temperature response applied throughout the troposphere (in the case of the Planck feedback), and the departure at each level from that uniform change (for the lapse rate feedback). We then integrate from the surface to the tropopause, defined as 100 mb at the equator and decreasing linearly to 300 mb at the poles.

Clouds are handled differently from non-cloud feedbacks, because the radiative effect of vertically-overlapping cloud fields is too nonlinear for the kernel method. Following Soden et al. (2008), the cloud feedback is calculated from the change in cloud radiative forcing (ΔCRF), with adjustments for cloud masking:

$$\lambda_c \Delta \overline{T}_s = \Delta CRF + (K_T^0 - K_T)dT + (K_q^0 - K_q)dq + (K_\alpha^0 - K_\alpha)d\alpha + (\Delta \widetilde{R}_f^0 - \Delta \widetilde{R}_f) \quad (4)$$

where K^0 terms are the clear-sky kernels, $\Delta \widetilde{R}_f^0$ is the clear-sky forcing, and ΔCRF is defined as the difference between net downward radiative fluxes in all-sky (i.e., the observed meteorological conditions, including clouds if present) and clear-sky (i.e., assuming no cloud) conditions. A discussion of the effect of clouds on clear-sky feedbacks can be found in Soden et al. (2004). As a consequence of this calculation, our nonlinear term in Equation 2 refers to clear-sky physics only (see Appendix A). Neglecting to account for the cloud-masking adjustments (e.g. Cess et al. 1990; Gregory and Webb 2008) may lead to misdiagnosis of the cloud feedback, as pointed out by Colman (2003). Note that the final term of the right-hand side of Equation 4 ensures that temperature-independent changes in clouds due to CO_2 forcing are not included in the cloud feedback.

In defining the control climate as the $1\times CO_2$ integration rather than as the $2\times CO_2$ fixed-SST integration, we run the risk of double-counting the temperature-independent response to CO_2, which has already been included with the forcing component $\Delta \widetilde{R}_f$. In essence, any change in climate field can be linearly related to surface temperature (as the feedback framework presumes), or not—in which case dx in Equation 3 or more likely ΔCRF in Equation 4 could include an additional source of nonlinear behavior. However as we will demonstrate, the two metrics of forcing (the difference between which *is* the semi-direct effect of CO_2) produce feedbacks that are not substantially different, lending confidence that the effect of the temperature-independent component is small within our kernel-computed feedbacks. Further, we identify the dominant source of the residual nonlinearity (see Section

3c) as due to a different process entirely.

3. Results

a. Feedbacks

Global-mean feedbacks are presented in Table 1. We first focus on the top row, which are the feedbacks calculated assuming the fixed-SST climate forcing. The temperature feedback is strongly negative (i.e., stabilizing the climate): A warmer planet emits more radiation to space (Planck feedback), and the weakened lapse rate, which is a consequence of moist adiabatic stratification, leads to emission from an even warmer atmosphere than if lapse rate were fixed (lapse rate feedback). The water vapor feedback is strongly positive because humidity is highly sensitive to warming, and because moistening the atmosphere increases infrared opacity and downwelling radiation. The surface albedo feedback is positive and, as expected, controlled by sea-ice processes. The net cloud feedback seems to be partly driven by changes in cloud fraction: the longwave cloud feedback is associated with the insulating effect of widespread increases in high cloud fraction, and the shortwave cloud feedback is associated with widespread decreases in reflective low cloud fraction (Fig. 4b). These global-mean feedbacks are in broad agreement with coupled-model studies, though our shortwave cloud feedback is on the high end of the range (e.g., Randall et al. 2007). Preliminary results indicate the absence of a tropical Walker circulation in the aquaplanet to be a controlling factor in the shortwave component of the cloud feedback, which may help to explain the relatively high sensitivity exhibited by our aquaplanet.

The sum of the linear feedbacks, which we call “total feedback” for convenience, is small and negative (-0.49 W m^{-2} K^{-1}). If the assumption of linearity were correct, then the global climate sensitivity would be calculated as $\Delta\overline{T}_s = \Delta\overline{\widetilde{R}_f}/\overline{\sum_x \lambda_x} = 7.7$ K, rather than the actual value of 4.69K. This points, then, to a substantial role for the nonlinear term. While it is smaller in magnitude than any individual feedback, comparison of the last two columns of Table 1 shows that the nonlinear term is 67% of the total feedback. Thus nonlinearities are of comparable importance to the linear feedbacks in affecting the TOA energy balance, at least in a global-mean sense and for this model set-up. Moreover this term tends to have a compensating role, in that it reduces global climate sensitivity. The importance of the nonlinearity in the global mean is further motivation to analyze the spatial pattern of the nonlinearity and feedbacks.

How does the magnitude of our nonlinearity compare to previous work? Though reporting conventions vary for the validity of the linear approximation, we can perform two crude comparisons. First, we estimate the equivalent nonlinear term from other studies by applying their cited values of feedbacks, forcing, and climate sensitivity to our Equation 2. Thus our nonlinear term, -0.33 W m^{-2} K^{-1}, is comparable in magnitude to estimates 0.39 W m^{-2} K^{-1} (Soden and Held 2006; Soden and Vecchi 2011, for GFDL CM2.1) and 0.13 W m^{-2} K^{-1} (Shell et al. 2008, for CAM3), though our sign is different. Second, as an alternative approach, we instead assume the nonlinear term can be expressed in the form $c\Delta\overline{T}_s^2$, such that the value of the coefficient c is a measure of the degree of nonlinearity. Roe and Armour (2011, their supplementary materials) report $|c| \leq 0.06$ W m^{-2} K^{-2} from a dozen different studies, with no consensus on sign. For our present study, the nonlinear term divided by $\Delta\overline{T}_s$ (-0.33 W m^{-2} K^{-1} per 4.69 K) gives $c = -0.07$ W m^{-2} K^{-2}. Thus the magnitude of our nonlinear term is roughly comparable to previous research, though on the high end. This may reflect our high climate sensitivity, or be a reflection of the idealized framework. That the nonlinear term is such a large percentage of the total linear feedback is a consequence of the total feedback being small.

For the sake of comparison, Table 1 also shows global-mean feedbacks for the stratosphere-adjusted radiative forcing. Due to the way in which the feedbacks are calculated, the choice of forcing can only affect the cloud feedback (compare Eqns. 3 and 4), total feedback, and residual. Overall, the differences in these terms as a function of forcing are fairly small. In fact, we find that the rapid tropospheric adjustment (included in the fixed-SST forcing) accounts for only a 16% decrease in the global shortwave cloud feedback, which is less than cited in previous studies (Colman and McAvaney 2011; Andrews et al. 2011). The discrepancy may reflect the inability of non-aquaplanet models to easily constrain land-temperature change, or alternately, a genuine difference in cloud response between models or model configurations. Hereafter we use only the fixed-SST forcing.

We now turn to the meridional structures of the feedbacks, which are shown in Figure 3. The first thing to note is that, converted to the same scale, the climate forcing has a value of about 0.5-1.2 W m^{-2} K^{-1} (2.5 to 5.5 W m^{-2} per 4.69 K). In other words, Figure 3 shows that the local adjustments by atmospheric process (i.e., feedbacks) are in general larger than the forcing itself. Another striking feature is that the Planck feedback is most strongly stabilizing (i.e., most negative) at high latitudes. This is in contrast to the simple picture one might naively expect from the Stephen-Boltzmann Law, wherein the change in outgoing flux varies as $4\sigma T^3$ and therefore is greatest in the tropics. However from Equation 3 we see that the Planck feedback is the product of the temperature kernel, $\partial R/\partial T$, whose amplitude indeed peaks at low latitudes (Fig. B1a in Appendix B), and the ratio $dT_s/d\overline{T}_s$. Given strong polar amplification (i.e., $dT_s \gg d\overline{T}_s$), this is enough to produce a Planck feedback that maximizes in magnitude at high latitudes. If feedbacks were instead defined as a Taylor series expansion around the local surface temperature change, as in Section 3b, then the pattern would be quite different (Feldl and Roe 2013). The lapse rate feedback is most negative where temperatures follow a moist adiabat (i.e., in the tropics) and most positive in the presence of high-latitude temperature inversions. The combined temperature feedback (Planck plus lapse rate, not shown) is strongly negative and peaks in magnitude at the equator.

The water vapor feedback is positive at all latitudes. However, the water vapor feedback is strongest where humidity is most sensitive to warming (c.f. Fig. B1b in Appendix B). These conditions occur in the subtropics and tropics, albeit the water vapor feedback is weaker along the equator due to high cloud masking of the tropical moistening at the ITCZ.

A key point here is that the water vapor feedback is not independent of the cloud fields, and this interaction between feedbacks hints at the presence of nonlinearity. In other words, water vapor changes under clouds have a reduced effect on the TOA fluxes, compared to cloud-free conditions. The water vapor feedback pattern is particularly sharp due to our perpetual equinox conditions (i.e., lack of seasonality) and aquaplanet configuration. In other words, we anticipate that the annual average over seasons would be smoother (i.e., exhibiting a less pronounced tropical minima) than the annual average over twelve months of a stationary ITCZ.

The net cloud feedback is positive everywhere except at high latitudes. The breakdown into shortwave and longwave components is shown in Figure 4. Changes in cloud fraction (Fig. 4b) are consistent with much of the meridional structure, though changes in cloud altitude and optical depth may also play a role (e.g., Colman et al. 2001; Zelinka et al. 2012). Recall that warming associated with a positive cloud feedback can occur by *decreases* in bright clouds (i.e., the SW effect) or *increases* in high, insulating clouds (i.e., the LW effect). The first thing to note from Figure 4a is that the shortwave component dominates the sign of the net response observed in Figure 3. Hence the peak in the net cloud feedback in the tropics is consistent with a decrease in cloud fraction at all levels, but especially in the upper troposphere (with some compensation between a positive shortwave and negative longwave cloud feedback); these cloud fraction changes are consistent with a weakening of the Hadley Cell. The negative net cloud feedback in the high latitudes coincides with an increase in low, bright clouds, and a poleward shift of the storm track. The positive net cloud feedback at intermediate, extratropical latitudes is consistent with widespread decreases in low cloud fraction (i.e., positive shortwave cloud feedback) and increases in high cloud fraction (i.e., positive longwave cloud feedback).

The surface albedo feedback is locally the strongest positive feedback, though it is confined to the vicinity of the ice line (Fig. 3). Consistent with expectations, reduction of sea-ice cover and the corresponding decrease in surface albedo in a warmer world lead to an increase in absorbed solar radiation, and further warming. Note that a compensation between positive albedo and negative shortwave cloud feedback is observed in Figures 3 and 4. This is a robust result across intermodel comparisons (Zelinka and Hartmann 2012; Crook et al. 2011), though the extent to which clouds are modified by increases in water vapor and evaporation over newly-open water is not easily constrained in a linear feedback framework (Bony et al. 2006; Stephens 2005). Previous studies have also pointed to an increase in high-latitude cloud optical depth due to increases in cloud water content, as well as phase changes (Senior and Mitchell 1993; Tsushima et al. 2006; Zelinka et al. 2012).

The meridional structure of the total feedback is the sum of the individual feedbacks, and is shown in Figure 3. Overall, the feedback is negative and stabilizing at high-latitudes (with the exception of the ice-line, where the albedo feedback is strong enough to result in a total feedback approaching zero). This locally negative total feedback might lead one to expect a weak surface temperature response. Yet Figure 1 shows strong polar amplification. Further, the total feedback is generally positive in the subtropics, which would imply a locally unstable climate—and an infinite response. Clearly then, either substantial redistribution of energy by meridional transport must occur, or else nonlinear interactions must arise. This finding is reminiscent of the work of Pierrehumbert (1995), in which circulation acts to shunt energy from unstable to stable latitudes, which are likened to "radiator fins." The general tendency of the total feedback to become more negative towards higher latitudes can also be seen in previous studies: although a large spread exists among models, Zelinka and Hartmann (2012) find that the zonal-mean total feedback parameter averaged over 12 CMIP3 models exhibits a tropical peak. It is not clear if their tropical peak (rather than our subtropical peak) is an artifact of the ensemble average, or if the absence of seasonality in our idealized framework accounts for the difference in location of the unstable domain. In any case, the combination of strong polar amplification and positive subtropical feedbacks implies critical roles for meridional transport and/or nonlinearities, to which we now turn.

The trade-off between meridional transport and the local demands of linear feedbacks is reflected in the three-term energy balance of Equation 2. The solid gray line in Figure 5 shows the meridional structure of the combined feedback and forcing term (i.e., $(\sum_x \lambda_x)\,\Delta\overline{T}_s + \Delta\widetilde{R}_f$). The positive values equatorward of approximately 40° represent a local warming tendency. In a perfectly linear world, the changes in transport (dashed line) would exactly balance the combined feedbacks and forcing. However in a nonlinear world, that adjustment is incomplete, and the remainder of the energy balance is accommodated by the nonlinear, or residual, term (solid black line in Fig. 5). In particular, there is increased meridional transport out of the subtropics, and the shape of this term closely mirrors that of the feedback-plus-forcing. In other words, in the subtropics, the system attempts to diverge heat away from the region of strong positive feedback, but transport alone does not fully accommodate this energy. The balance is taken up by the nonlinear term, which provides a cooling tendency in the low latitudes (equatorward of 50°) and a warming tendency elsewhere. Hence in addition to compensating the global sensitivity, the nonlinear term plays an important, compensating role at many latitudes: It opposes the positive feedback in the tropics, and likewise offsets the negative feedback at high latitudes. Further, the nonlinear term is minimized (i.e., the assumption of linearity works best) in the midlatitudes; a negative total feedback is balanced by anomalous heat convergence at 45°. Our ability to assess the nonlinear contribution is a key strength of our approach.

b. *Polar amplification*

Polar amplification is a striking feature of all climate models predictions and is also observed in global temperature trends (Solomon et al. 2007). In our simulation we see two scales to the polar amplification: an enhancement of the temperature response polewards of about 30°, and a much larger enhancement polewards of 60°. Polewards of 60°, the average warming is 2.2 times the global-mean response; this degree of amplification is consistent with other studies (Hwang et al. 2011; Holland and Bitz 2003). We can apply the feedback framework toward understanding polar amplification in terms of the spatial patterns of climate feedbacks, forcing, heat transport, and nonlinearities. In what follows, we pursue the apparently contradictory result that the temperature response is largest in regions where the feedback is most stabilizing (compare Figs. 1 and 3).

Equation 2 can be rewritten with local temperature change ΔT_s substituted for global-mean $\Delta \overline{T}_s$, and the Planck feedback λ_P separated from the non-Planck feedbacks $\sum \lambda_{NP}$:

$$\Delta T_s = \frac{1}{\lambda_P}\left[\Delta R - \left(\sum_i \lambda_{NP_i}\right)\Delta T_s - \Delta \widetilde{R}_f - \mathcal{R}\right]. \tag{5}$$

In essence, we normalize the terms in the energy balance by the Planck feedback. This weighting avoids an undefined surface temperature response where the total feedback goes to zero. The feedback term in Equation 5 is also more similar in form to the conventional definition where feedback factor $f = -\lambda_{NP}/\lambda_P$ (e.g., Roe 2009). Thus the pattern of local temperature response is given as the partial temperature change attributed to each term on the right-hand side of Equation 5; this decomposition is also utilized by Crook et al. (2011). These individual contributions as a function of latitude are presented in Figure 6, together with the total surface temperature change, shown in grey. As a reminder, the non-Planck feedbacks include lapse rate, water vapor, surface albedo, and cloud feedbacks.

The forcing produces a small and uniform warming of 0.9-1.6 K (red line, Fig. 6). This contribution to surface temperature change is not substantially different when the stratosphere-adjusted forcing is instead used (not shown). In other words, the previously noted asymmetries in forcing are small compared to the other terms in affecting surface temperature. The nonlinear term is also small (green line, $\pm$1.6 K) and, as expected from Figure 5, cools the tropics and warms the high latitudes, contributing to the polar-amplified shape of the warming pattern. On average the transport term also exhibits a pattern of tropical cooling and high-latitude warming, consistent with a poleward export of heat from the tropics, though its meridional structure and magnitude are more variable. The non-Planck feedbacks provide a warming tendency at all latitudes, and are the major contributor to the more than 10 K warming near the ice-line. In general, non-Planck feedbacks and transport exhibit strong compensation, while the nonlinear term and forcing make smaller contributions to surface temperature change and with less meridional variability. Overall then, the enhancement of the *average* response poleward of 30°, relative to the response equatorward of 30°, may be attributed predominantly to the change in sign of the transport term (and to a lesser degree, the nonlinear term). The pole-to-equator *shape* of the polar amplification is largely explained by the combined effects of feedbacks and transport.

The further amplification of surface temperature polewards of 60° may be characterized in two parts: non-Planck feedbacks (particularly surface albedo, longwave cloud, and lapse rate feedbacks, see Figs. 3 and 4) from 60-70°, and meridional heat transport of 4.7 K poleward of 70°. The strong warming tendency of the non-Planck feedbacks at the ice-line is partially offset by the transport term (i.e., a cooling tendency due to heat export). Polewards of the ice-line there is anomalous convergence of at least a portion of this exported heat, which maintains the enhanced warming right to the poles. At the poles, none of the terms act as cooling tendencies. Hence we find a consistent picture at both hemispheric and regional scales, in which local temperature change is controlled by anomalous heat divergence away from regions of strong positive feedbacks (i.e., the ice-line and the subtropics) and convergence into regions of more negative feedbacks (i.e., the midlatitudes and poles).

The influence of the Planck weighting in Equation 5 is demonstrated in the lower panel of Figure 6. The dashed line shows how the predicted surface warming would change if the global-mean weighting $\overline{\lambda_P^{-1}}$ had been used in Equation 5, instead of the full spatial field. The meridional structure of the Planck feedback, which increases in magnitude toward the poles (see Fig. 3), contributes an additional 23% warming in the high latitudes (poleward of 60°) and 15% cooling in the subtropics (5-25°). Thus the Planck feedback comes in at tertiary importance, behind the other feedbacks and transport, in explaining polar amplification, though its approximately 2 K high-latitude warming is distributed amongst the other terms and cannot be easily isolated.

Our results have demonstrated the importance of meridional heat fluxes to the system response. We next consider the breakdown of the transport term into changes in latent and dry-static energy flux, following Trenberth and Stepaniak (2003) and Hwang and Frierson (2010). As part of the calculation, we subtract the surface flux from the TOA flux, in order to solve for the total atmospheric (i.e. moist-static energy) budget; the surface flux includes contributions from net downward radiation at the surface, sensible heat flux, and latent heat flux due to evaporation and melting snowfall into the ocean. We find the change in surface flux to be smaller than ± 0.73 W m^{-2} at all latitudes and negligible in the global mean. The northward latent energy flux is calculated as the integral, with respect to latitude, of evaporation minus precipitation (multiplied by the latent heat of vaporization for consistent units), and the dry-static energy flux is then the residual of the latent and total atmospheric fluxes.

Changes in northward energy fluxes are shown in Figure 7. Positive slopes in the figure correspond to regions of anomalous flux divergence, and negative slopes to anomalous convergence. The total flux change (gray line) confirms an increase in divergence away from the subtropics, and an anomalous divergence away from the ice line (i.e., decreased convergence with respect to the control climate). Relative to the total flux change, the latent and dry-static energy components are large and mostly compensating. In the warmer climate, there is an increase in latent energy flux poleward of approximately 25-30° (solid black line). This is significantly offset by a decrease in dry-static energy flux (dashed line), presumably due to weaker midlatitude temperature gradients. However the total flux change is still positive, and thus it is the *larger* increase in latent energy flux that explains the contribution of transport to polar amplification polewards of 30°. Interestingly, the dry-static energy gradient weakens considerably polewards of the ice line. Therefore the contribution of heat transport to polar amplification at the highest latitudes (see also Hwang et al. 2011; Langen et al. 2012) is driven solely by the latent energy flux convergence, with no compensation from dry-static energy. Figure 7 also shows an increase in equatorward latent heat flux and an increase in poleward dry-static energy flux, which have the same sign as the climatological fluxes.

c. *Source of the nonlinearity*

Up to this point, we have characterized the residual nonlinearity, without addressing which interactions between feedbacks are responsible for the term. The core of the issue is that the kernel framework assumes each variable and each vertical level are independent and can be linearly combined. Whereas in fact, vertical masking of clear-sky variables, and interactions amongst these variables, could complicate this picture. Analogously, it is well known that clouds mask underlying tropospheric changes (Soden et al. 2004). Water vapor exhibits similar behavior. Figure 8a shows changes in specific humidity between the $1\times CO_2$ and $2\times CO_2$ experiments. These changes show an overall moistening and are consistent with a weakening and expansion of the Hadley Cell (e.g., Held and Soden 2006). The linear model (e.g. kernel approach) assumes that changes in mid- and lower-tropospheric water vapor have as large of an effect on the TOA as changes aloft. In actuality, the sensitivity of TOA radiation fluxes to upper tropospheric humidity is well known (Cess 1975; Spencer and Braswell 1997), and we expect the TOA balance to be most affected by changes aloft. As a result, we anticipate that the water vapor kernel would result in an overestimate of the TOA fluxes in regions of strong upper-level moistening, which would manifest as a negative nonlinearity.

We test this hypothesis by running the actual changes at all levels ($2\times CO_2$ minus $1\times CO_2$) in humidity, temperature, and surface albedo *simultaneously* through the offline radiation code to calculate the magnitude of the TOA fluxes. The net radiative flux at the TOA is then compared to the linear sum of the individual variables at each level, which is what the kernel framework presumes. Results are shown in Figure 8b. The solid line can be thought of as the difference between the GCM response and the linear approximation, or in other words, it represents an independent measure of the nonlinearity. We see that this difference does a rather remarkable job of capturing the magnitude and qualitative shape of the residual nonlinearity (dashed line), with some obvious departures.

Remaining sources of nonlinearity (i.e., the difference between the two lines on Figure 8), can be considered with the help of Equation A2. We have already accounted for nonlinearities within the third term on the right-hand-side, $\sum_n \lambda_n^0 \Delta \overline{T}_s$, i.e., vertical masking of, and interactions between, clear-sky feedbacks. As mentioned in Section 2b, there is also the possibility of double-counting the rapid tropospheric adjustment to CO_2. However we expect this contribution to be minor because the residual is nearly identical when calculated with a stratosphere-adjusted, rather than fixed-SST, radiative forcing (not shown), which does not suffer from double counting. Hence any remaining nonlinearities may be attributed to second-order terms associated with the effect of clouds on non-cloud fields. First-order terms were accounted for in Equation 4, following advances by Soden et al. (2008), and it is straightforward to show that a quadratic form of Equation 4 would propagate additional terms to Equation A2.

4. Summary and Discussion

In this study we have sought to understand the spatial structure of climate feedbacks and the relative importance of nonlinearities and meridional heat transport. We have designed a clean experiment, which seeks to remove as many of the common energy-balance approximations as possible. In particular, we employ a simplified aquaplanet model, and explicitly calculate both fixed-SST climate forcing and radiative kernels for this precise set-up. Our high climate sensitivity of 4.69 K is consistent with large subtropical regions of positive water vapor and cloud feedbacks. Two regions of positive feedbacks, the subtropics and the ice-line, force anomalous divergence of heat flux, which translates into polar amplification of the surface temperature response via meridional latent heat transport. Nonlinearities reinforce this pattern of tropical cooling and high-latitude warming tendencies, and also reduce global climate sensitivity from very high to merely high. The nonlinear term can be thought of as reinforcing the transport-induced warming, or, alternatively, as offsetting the total linear feedback. The resulting polar-amplified warming bears the signature of feedbacks, transport, and nonlinearities, but importantly, is not limited to the latitude where a particular physical process is active.

One of the goals of this research has been to understand how local processes affect non-local climate responses. The feedback pattern is characterized by strongly positive subtropical feedbacks, and the temperature response pattern characterized by polar amplification. Clearly, meridional heat transport matters for redistributing energy. Indeed we find that transport plays a role on a couple of spatial scales—from the subtropics to the mid-latitudes, and from the high-latitudes, poleward. The stable midlatitudes also display interesting complexity: abutting regions of positive feedbacks contribute to a maximum increase in heat convergence at 45°, near the latitude where the nonlinearity is minimized.

We have further studied the source of our "nonlinear" term, which strictly represents the clear-sky residual between the energy-flux changes predicted by linear theory and the actual, model-produced flux changes. Though a modest contributor of at most 2K to local temperature response (when normalized by the Planck feedback), the meridional structure of the nonlinearity and its tendency to compensate climate feedbacks suggest a physical mechanism at work. Indeed from a Taylor-series perspective, these nonlinearities can be thought of as higher-order terms that do not scale linearly with surface temperature change (e.g., Stephen-Boltzmann Law, or Clausius-Clapeyron relationship) or interactions between feedbacks (i.e., cross-terms in the energy budget). The effect is such that, at low latitudes, the feedback is less than the sum of its parts and at high latitudes it is more than the sum of its parts. Generally speaking, our results caution against the use of methods in which the residual is subsumed into one of the linear feedbacks (e.g., the cloud feedback of Soden and Held 2006).

Through offline radiation experiments, we have attributed the bulk of the nonlinear term to interactions within (i.e., vertical masking) and amongst clear-sky feedbacks, and pointed to quadratic effects of clouds on non-cloud variables as the leading candidate for remaining nonlinearity. Hence nonlinear feedbacks may represent dynamical constraints within the system: changes in atmospheric circulation modulate the degree of compensation between terms in the energy balance, in a way not accounted for by the linear, Taylor-series approximation. For example, dynamically controlled changes in specific humidity were suggested to be a major cause of the nonlinearity. Nonlinear interactions amongst temperature, humidity, and surface albedo, with the latter constrained to high latitudes, are active as well. It should be noted that the wholesale substitution of climate variables, all levels at once, is a feature of both our offline radiation experiments and the PRP method of calculating feedbacks. Thus we would expect that nonlinearities arising from the PRP method to be restricted to interactions amongst (but not within, as in vertical masking) feedbacks.

The idealized aquaplanet framework provides a unique lens on radiative interactions in a changing climate, though some of our results may be a consequence of experimental design. We are confident the residual does represent an approximation of the nonlinearity, because (1) we made every effort to close the energy balance as nearly as possible, by diagnosing radiative kernels and forcing for this model setup, and (2) two independent estimates of the nonlinear term (residual and offline calculations) are consistent. However the aquaplanet simulation is, by its very nature, simplified. For instance, lack of land-sea contrast will have a profound effect on cloud climatologies, which we have mentioned with respect to our shortwave cloud feedback—though this perhaps matters less to the nonlinearity, which is, of course, a clear-sky effect. While it is reassuring that our global-mean feedbacks are within the spread of intermodel comparisons (Bony et al. 2006; Randall et al. 2007), future work will systematically relax the simplifying assumptions towards greater realism.

Possible avenues of progress include the following: (1) a comparison of the kernels from $1\times CO_2$ and $2\times CO_2$ climatologies, in order to address the mean-state dependence, or extension of the method to calculate second-order terms; (2) the inclusion of greater realism such as a seasonal cycle or ocean heat transport, which would directly affect the energy balance via the transport term and indirectly through the coupling between feedbacks and surface response; and (3) use of a wider range of forcings to address how feedbacks behave for larger surface-temperature changes (e.g., Colman and McAvaney 2009). Furthermore, the equilibrated climate change must satisfy both radiative and dynamical constraints. Sharply honed numerical experiments that address the conditions under which either dynamics or radiation dominates the response would be useful.

Our breakdown of the meridional structure of temperature response into individual components (Equation 5) also illustrates some issues for the predictability of regional climate change. Local feedbacks alone do not set the pattern of temperature response: atmosphere (and ocean) dynamics act to redistribute energy in the system, and so one must constrain the feedbacks everywhere in order to constrain the response anywhere. Figure 6 shows the partial temperature change for feedbacks, transport, forcing, and nonlinearities as a function of latitude in our simulations. It also provides some sense for how the meridional structure of predicted climate change might vary, if improved understanding resulted in a different pattern of total feedback.

Conventional climate feedback analysis characterizes only the energy balance and is inherently linear by construction. We have extended that perspective in an idealized framework to include nonlinear terms and to consider nonlocal effects. These must operate in the real climate system and are an important component of understanding predictability. The meridional structures of individual feedbacks are governed by the classical climatic zones (i.e., the ITCZ, the subtropics, the midlatitudes, the poles), and thus are a consequence of mean-state dynamics. However dynamical changes in the circulation pattern may modulate nonlinearities and, as a consequence, global climate sensitivity. Further, the system tends to allocate energy towards latitudes that can most effectively radiate to space. This means that warming is minimized in the subtropics in spite of strong positive feedbacks. A complete picture of climate sensitivity must unify dynamical and radiative frameworks, and it is our hope that the current study offers some insights into what that may entail.

Acknowledgments.

The authors thank Dargan Frierson, Mark Zelinka, Ben Sanderson, Dennis Hartmann, and Kyle Armour for valuable discussions; Marc Michelsen for technical assistance; and two anonymous reviewers and the editor. This work was partially supported by the National Science Foundation (EAR-0908558).

APPENDIX A

Why a clear-sky residual?

The clear-sky (rather than all-sky) residual is a consequence of our cloud feedback calculation. Equation 2 can be rearranged to give

$$\Delta R = \Delta \widetilde{R}_f + \left(\sum_n \lambda_n \right) \Delta \overline{T}_s + \lambda_c \Delta \overline{T}_s, \tag{A1}$$

where the cloud feedback λ_c is split from the other, non-cloud feedbacks ($n = T, q, \alpha$). Substituting Equation 4 into Eq. A1 gives

$$\Delta R = \Delta \widetilde{R}_f^0 + \Delta CRF + \left(\sum_n \lambda_n^0 \right) \Delta \overline{T}_s, \tag{A2}$$

where superscripted terms represent clear-sky fluxes. Hence the residual becomes

$$\mathcal{R} = (\Delta R - \Delta CRF) - \left[\Delta \widetilde{R}_f^0 + \left(\sum_n \lambda_n^0 \right) \Delta \overline{T}_s \right], \tag{A3}$$

or the difference between actual, model-produced clear-sky fluxes ($\Delta R - \Delta CRF$) and kernel-approximated clear-sky fluxes (the remaining terms).

APPENDIX B

Radiative kernels

To facilitate comparison with previous studies (Soden and Held 2006; Soden et al. 2008; Shell et al. 2008), we present height-latitude cross sections of our perpetual equinox, aqua-planet kernels. The kernels in Figure B1 represent the contribution of each level and latitude to the change in longwave TOA fluxes. The temperature kernel (Fig. B1a) is strongly negative (i.e., stabilizing the climate) because an increase in temperature increases OLR, following the Stefan-Boltzman Law. Under clear skies (not shown) the sensitivity peaks in the tropics where temperatures are highest. However all-sky TOA fluxes are sensitive to cloud-top temperature, with the largest contributions from regions of high convective clouds and subtropical and midlatitude boundary layer clouds. The surface component of the temperature kernel (Fig. B1d) exhibits cloud masking, with decreased sensitivity aligned beneath regions of high cloudiness. Cloud-masking effects are also apparent in the surface albedo kernel (Figure B1c), though this kernel obviously only matters near the climatological ice-line.

The water vapor kernel (Fig. B1b) shows the TOA radiative flux response to atmospheric moistening. In calculating the kernel, specific humidity q was perturbed to match the change in saturation specific humidity that would occur from a 1 K warming, assuming fixed relative humidity (Soden and Held 2006). Positive values indicate that an increase in atmospheric water vapor leads to an increase in infrared opacity and downwelling radiation (decreasing OLR), consistent with the role of water vapor as a greenhouse gas. High sensitivity in the tropics is also influenced by self-broadening of water vapor absorption spectra (Shine and Sinha 1991). At high latitudes and low levels, the water vapor kernel is negative (an anti-greenhouse effect); the effect of humidifying the atmosphere is to raise the emission level (Cess 1975; Held and Soden 2000), leading to an increase in OLR in regions of temperature inversions. The water vapor kernel peaks strongly in the climatologically dry upper troposphere because of the high sensitivity of saturation vapor pressure at very cold temperatures and low pressures (via the Clausius-Clapeyron relationship); for fixed relative humidity at 200 K, specific humidity changes by 15%/K (Held and Soden 2000). Hence the pattern of this kernel is tied to the assumption of fixed relative humidity. If relative humidity were instead allowed to decrease, then warming would not require moistening, and it would be possible to imagine a weakened water vapor response in the subtropics—though the lapse rate would adjust accordingly to compensate this effect (Bony et al. 2006).

REFERENCES

Andrews, T. and P. M. Forster, 2008: CO_2 forcing induces semi-direct effects with consequences for climate feedback interpretations. *Geophys. Res. Lett.*, **35 (4)**, doi: 10.1029/2007GL032273.

Andrews, T., J. M. Gregory, P. M. Forster, and M. J. Webb, 2011: Cloud adjustment and its role in CO_2 radiative forcing and climate sensitivity: A review. *Surveys in Geophysics*, doi:10.1007/s10712-011-9152-0.

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. *Geophysical Research Letters*, **39 (9)**, doi:10.1029/2012GL051607.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2012: Time-varying climate sensitivity from regional feedbacks, submitted to Journal of Climate.

Bode, H. W., 1945: *Network Analysis and Feedback Amplifier Design*. Bell Telephone Labs Series, Van Nostrand Reinhold, New York, 551 pp.

Bony, S., et al., 2006: How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, **19 (15)**, 3445–3482, doi:10.1175/JCLI3819.1.

Cess, R. D., et al., 1990: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *Journal of Geophysical Research*, **95 (D10)**, 16,601–16,615.

Cess, R. D., 1975: Global climate change: an investigation of atmospheric feedback mechanisms. *Tellus*, **27 (3)**, 193–198, doi:10.1111/j.2153-3490.1975.tb01672.x.

Colman, R. A. and B. McAvaney, 2011: On tropospheric adjustment to forcing and climate feedbacks. *Climate Dynamics*, **36 (9)**, 1649–1658.

Colman, R. A., S. B. Power, and B. J. McAvaney, 1997: Non-linear climate feedback analysis in an atmospheric general circulation model. *Climate Dynamics*, **13**, 717–731, 10.1007/s003820050193.

Colman, R. A., 2003: A comparison of climate feedbacks in general circulation models. *Climate Dynamics*, **20**, 865–873, doi:10.1007/s00382-003-0310-z.

Colman, R., J. Fraser, and L. Rotstayn, 2001: Climate feedbacks in a general circulation model incorporating prognostic clouds. *Climate Dynamics*, **18**, 103–122, doi: 10.1007/s003820100162.

Colman, R. and B. McAvaney, 2009: Climate feedbacks under a very broad range of forcing. *Geophysical Research Letters*, **36 (1)**, doi:10.1029/2008GL036268.

Crook, J. A., P. M. Forster, and N. Stuber, 2011: Spatial patterns of modeled climate feedback and contributions to temperature response and polar amplification. *Journal of Climate*, **24 (14)**, 3575–3592, doi:10.1175/2011JCLI3863.1.

Feldl, N. and G. H. Roe, 2013: Five perspectives on feedbacks, in preparation.

GFDL Global Atmospheric Model Development Team, 2004: The new GFDL Global Atmosphere and Land Model AM2–LM2: Evaluation with prescribed SST simulations. *Journal of Climate*, **17 (24)**, 4641–4673, doi:10.1175/JCLI-3223.1.

Gregory, J. M., et al., 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31 (3)**, doi:10.1029/2003GL018747.

Gregory, J. and M. Webb, 2008: Tropospheric adjustment induces a cloud component in CO_2 forcing. *Journal of Climate*, **21 (1)**, 58–71, doi:10.1175/2007JCLI1834.1.

Hansen, J., et al., 2005: Efficacy of climate forcings. *Journal of Geophysical Research*, **110 (D18)**.

Held, I. M. and K. M. Shell, 2012: Using relative humidity as a state variable in climate feedback analysis. *Journal of Climate*, **25 (8)**, 2578–2582, doi:10.1175/JCLI-D-11-00721.1.

Held, I. M. and B. J. Soden, 2000: Water vapor feedback and global warming. *Ann. Rev. Energy Environ.*, **25**, 441–475.

Held, I. M. and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *Journal of Climate*, **19 (21)**, 5686–5699, doi:10.1175/JCLI3990.1.

Holland, M. and C. Bitz, 2003: Polar amplification of climate change in coupled models. *Climate Dynamics*, **21**, 221–232.

Hwang, Y.-T., D. M. W. Frierson, and J. E. Kay, 2011: Coupling between arctic feedbacks and changes in poleward energy transport. *Geophysical Research Letters*, **38 (17)**, doi:10.1029/2011GL048546.

Hwang, Y.-T. and D. M. W. Frierson, 2010: Increasing atmospheric poleward energy transport with global warming. *Geophys. Res. Lett.*, **37 (24)**, doi:10.1029/2010GL045440.

Ingram, W., 2012: A new way of quantifying GCM water vapour feedback. *Climate Dynamics*, **Feb**, 10.1007/s00382-012-1294-3.

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. *Journal of Climate*, doi:10.1175/JCLI-D-11-00524.1.

Langen, P. L., R. G. Graversen, and T. Mauritsen, 2012: Separation of contributions from radiative feedbacks to polar amplification on an aquaplanet. *Journal of Climate*, **25 (8)**, 3010–3024, doi:10.1175/JCLI-D-11-00246.1.

Myhre, G., E. J. Highwood, K. P. Shine, and F. Stordal, 1998: New estimates of radiative forcing due to well mixed greenhouse gases. *Geophysical Research Letters*, **25 (14)**, 2715–2718.

Pierrehumbert, R. T., 1995: Thermostats, radiator fins, and the local runaway greenhouse. *Journal of the Atmospheric Sciences*, **52 (10)**, 1784–1806.

Ramaswamy, V., et al., 2001: Radiative forcing of climate change. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson, Ed., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881.

Randall, D., et al., 2007: Cilmate models and their evaluation. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Roe, G. H. and K. C. Armour, 2011: How sensitive is climate sensitivity? *Geophys. Res. Lett.*, **38 (14)**, doi:10.1029/2011GL047913.

Roe, G., 2009: Feedbacks, timescales, and seeing red. *Annual Review of Earth and Planetary Sciences*, **37**, 93–115.

Schlesinger, M. E., 1985: Feedback analysis of results from energy balance and radiative-convective models. *The Potential Climatic Effects of Increasing Carbon Dioxide*, M. C. MacCracken and F. M. Luther, Eds., U. S. Department of Energy, Washington, DC, DOE/ER-0237, 280–319.

Senior, C. A. and J. F. B. Mitchell, 1993: Carbon dioxide and climate: The impact of cloud parameterization. *Journal of Climate*, **6**, 393–418.

Senior, C. A. and J. F. B. Mitchell, 2000: The time dependence of climate sensitivity. *Geophysical Research Letters*, **27 (17)**, 2685–2688, doi:10.1029/2000GL011373.

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. *Journal of Climate*, **21 (10)**, 2269–2282, doi:10.1175/2007JCLI2044.1.

Shine, K. P., J. Cook, E. J. Highwood, and M. M. Joshi, 2003: An alternative to radiative forcing for estimating the relative importance of climate change mechanisms. *Geophys. Res. Lett.*, **30 (20)**, doi:10.1029/2003GL018141.

Shine, K. P. and P. M. Forster, 1999: The effect of human activity on radiative forcing of climate change: a review of recent developments. *Global and Planetary Change*, **20 (4)**, 205 – 225, doi:10.1016/S0921-8181(99)00017-X.

Shine, K. P. and A. Sinha, 1991: Sensitivity of the earth's climate to height-dependent changes in the water vapour mixing ratio. *Nature*, **354 (6352)**, 382–384.

Soden, B. J., A. J. Broccoli, and R. S. Hemler, 2004: On the use of cloud forcing to estimate cloud feedback. *Journal of Climate*, **17 (19)**, 3661–3665.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, **21 (14)**, 3504–3520, doi:10.1175/2007JCLI2110.1.

Soden, B. J. and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, **19 (14)**, 3354–3360, doi:10.1175/JCLI3799.1.

Soden, B. J. and G. A. Vecchi, 2011: The vertical distribution of cloud feedback in coupled ocean-atmosphere models. *Geophysical Research Letters*, **38**, doi:10.1029/2011GL047632.

Solomon, S., et al., 2007: Technical Summary. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the*

Intergovernmental Panel on Climate Change, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Spencer, R. W. and W. D. Braswell, 1997: How dry is the tropical free troposphere? Implications for global warming theory. *Bulletin of the American Meteorological Society*, **78 (6)**.

Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *Journal of Climate*, **18 (2)**, 237–273, doi:10.1175/JCLI-3243.1.

Trenberth, K. E. and D. P. Stepaniak, 2003: Covariability of components of poleward atmospheric energy transports on seasonal and interannual timescales. *Journal of Climate*, **16 (22)**, 3691–3705.

Tsushima, Y., et al., 2006: Importance of the mixed-phase cloud distribution in the control climate for assessing the response of clouds to carbon dioxide increase: A multi-model study. *Climate Dynamics*, **27**, 113–126, 10.1007/s00382-006-0127-7.

Watanabe, M., H. Shiogama, M. Yoshimori, T. Ogura, T. Yokohata, H. Okamoto, S. Emori, and M. Kimoto, 2011: Fast and slow timescales in the tropical low-cloud response to increasing CO_2 in two climate models. *Climate Dynamics*, 1–15, doi:10.1007/s00382-011-1178-y.

Wetherald, R. T. and S. Manabe, 1988: Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, **45 (8)**, 1397–1416.

Wyant, M. C., C. S. Bretherton, P. N. Blossey, and M. Khairoutdinov, 2012: Fast cloud adjustment to increasing CO_2 in a superparameterized climate model. *Journal of Advances in Modeling Earth Systems*, **4**.

Yoshimori, M., J. C. Hargreaves, J. D. Annan, T. Yokohata, and A. Abe-Ouchi, 2011:

Dependency of feedbacks on forcing and climate state in physics parameter ensembles. *Journal of Climate*, **24**, 6440–6455, doi:10.1175/2011JCLI3954.1.

Zelinka, M. D. and D. L. Hartmann, 2012: Climate feedbacks and their implications for poleward energy flux changes in a warming climate. *Journal of Climate*, **25**, doi:10.1175/JCLI-D-11-00096.1.

Zelinka, M. D., S. A. Klein, and D. L. Hartmann, 2012: Computing and partitioning cloud feedbacks using cloud property histograms. Part II: Attribution to changes in cloud amount, altitude, and optical depth. *Journal of Climate*, doi:10.1175/JCLI-D-11-00249.1.

List of Tables

Table 1: Global-mean, annual-mean feedbacks for fixed-SST (top row) and stratosphere-adjusted (bottom row) radiative forcings. Planck (P), lapse rate (LR), water vapor (WV), and albedo (A) feedbacks are unchanged as a function of forcing. The "total feedback" is the sum of the linear feedbacks. We interpret the residual as the nonlinear term. The terms in Equation 2 are normalized by the global-mean surface temperature change, such that units are given in W m^{-2} K^{-1} unless otherwise noted.

Forcing	Feedbacks								
(W m^{-2})	P	LR	WV	A	Net Cloud	LW Cloud	SW Cloud	Total	Residual
3.79	-3.03	-0.69	1.62	0.35	1.27	0.56	0.70	-0.49	-0.33
3.41	"	"	"	"	1.33	0.49	0.83	-0.43	-0.31

List of Figures

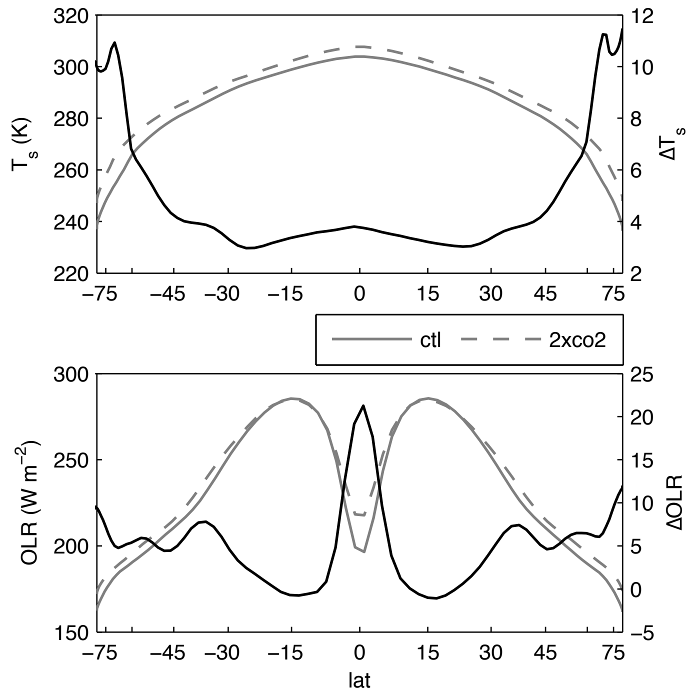

Figure 1: Zonal-mean, annual-mean T_s (top; K) and OLR (bottom; W m^{-2}) 10-year climatologies (gray) and anomalies (black). In both panels, solid gray lines indicate the $1\times CO_2$ climate; dashed lines, the $2\times CO_2$ climate. The global-mean equilibrium climate sensitivity is 4.69 K, though the meridional structure is strongly characterized by polar amplification.

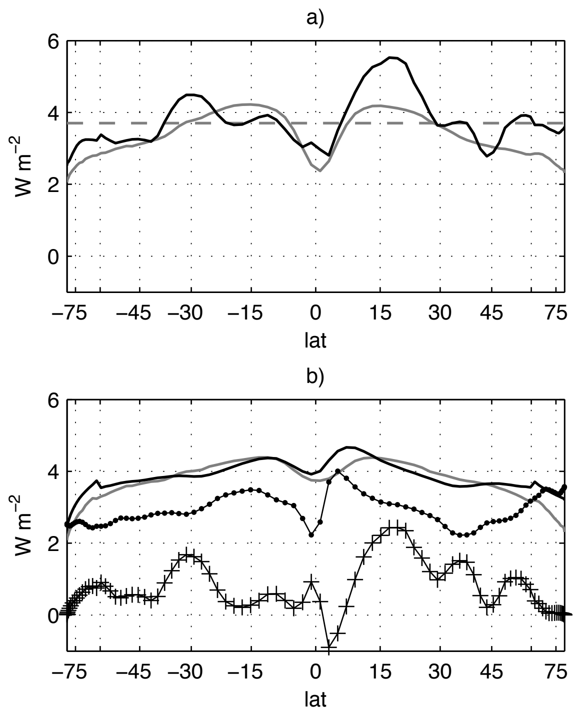

Figure 2: (a) Zonal-mean radiative forcing (W m^{-2}) for CO_2 doubling: uniform 3.7 W m^{-2} (dashed gray) from Myhre et al. (1998), which serves as the basis for the IPCC Third Assessment Report estimates; stratosphere-adjusted forcing calculated from the GFDL radiative transfer code (solid gray, averaged over two months of 8× daily model output); and fixed-SST forcing (solid black, averaged over 40 years), which includes rapid tropospheric adjustments. (b) LW (dotted) and SW (hashed) components of fixed-SST forcing. Net clear-sky stratosphere-adjusted forcing (solid gray) also shown for comparison to net clear-sky fixed-SST forcing (solid black).

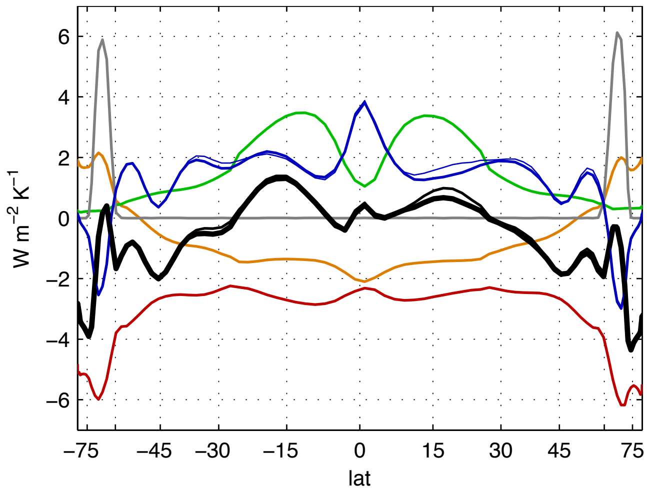

Figure 3: Zonal-mean, annual-mean feedbacks (W m^{-2} K^{-1}) for Planck (red), lapse rate (orange), water vapor (green), surface albedo (gray), cloud (blue), and the sum of these linear feedbacks (black). Thin lines represent feedbacks calculated with the stratosphere-adjusted, rather than fixed-SST, forcing.

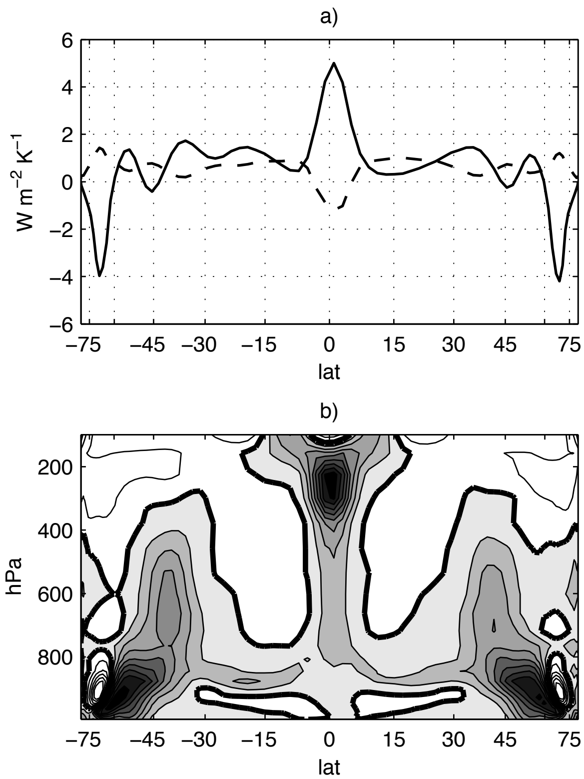

Figure 4: (a) Zonal-mean, annual-mean shortwave (solid) and longwave (dashed) components of the cloud feedback (W m^{-2} K^{-1}). (b) Change in cloud fraction. The zero contour is indicated by the heavy black line, and the contour interval is 2%; dark colors represent a decrease.

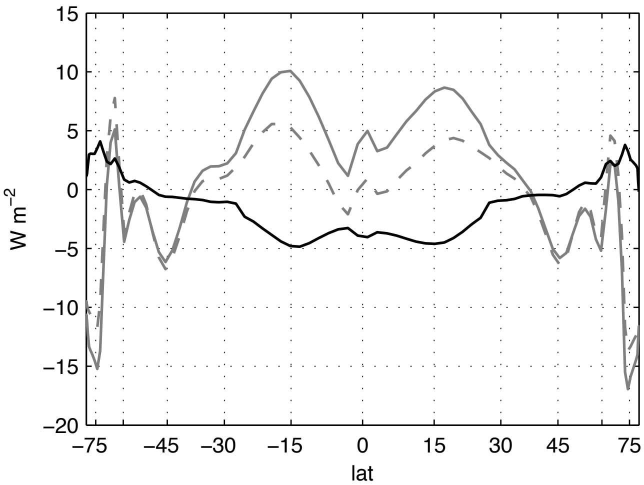

Figure 5: The balance of the three terms in Equation 2 (W m^{-2}). Recall that transport is the change in TOA net radiative flux, which in equilibrium must be equal to the change in convergence of atmospheric heat transport (i.e., $\Delta R = \Delta(\nabla \cdot F)$). The nonlinear term (black line) is calculated as the residual between meridional transport (dashed gray line) and the combined feedbacks and forcing (solid gray line).

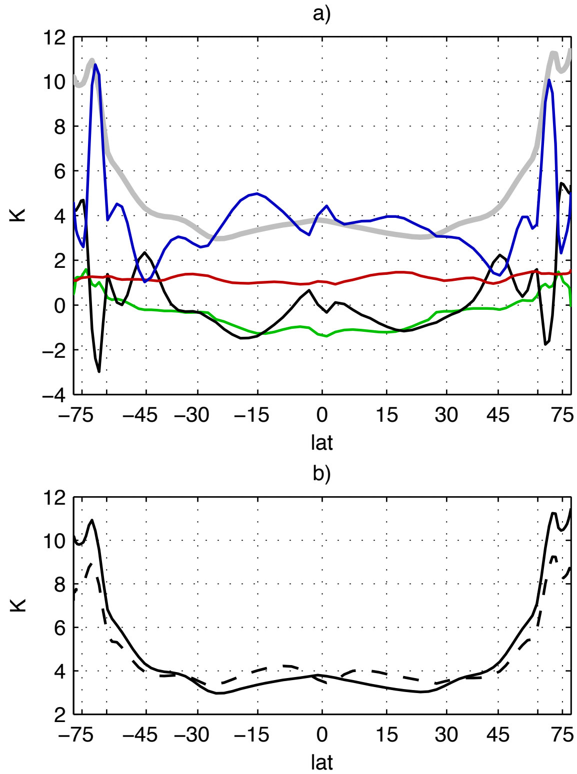

Figure 6: (a) Zonal-mean, annual-mean partial temperature changes (K) attributed to forcing (red), nonlinear term (green), sum of non-Planck feedbacks (blue), and transport (black). The total surface temperature change is shown in gray. Components are weighted by the Planck feedback, which has meridional structure. (b) Local temperature change ΔT_s (K) if global-mean weighting $\overline{\lambda_P^{-1}}$ were instead applied in Equation 5 (dashed line). Solid line reproduced from gray line in upper panel.

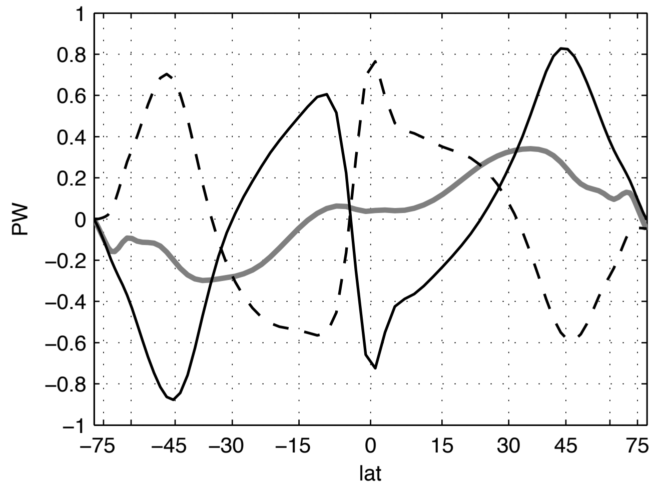

Figure 7: Zonal-mean, annual-mean change in northward energy flux (PW). The total northward energy flux (thick gray) is obtained by integrating with respect to latitude the sum of the TOA and surface fluxes. The latent energy (solid black) is calculated from the integrated evaporation minus precipitation, and and the dry static energy (dashed black) is from the residual of the other two fluxes.

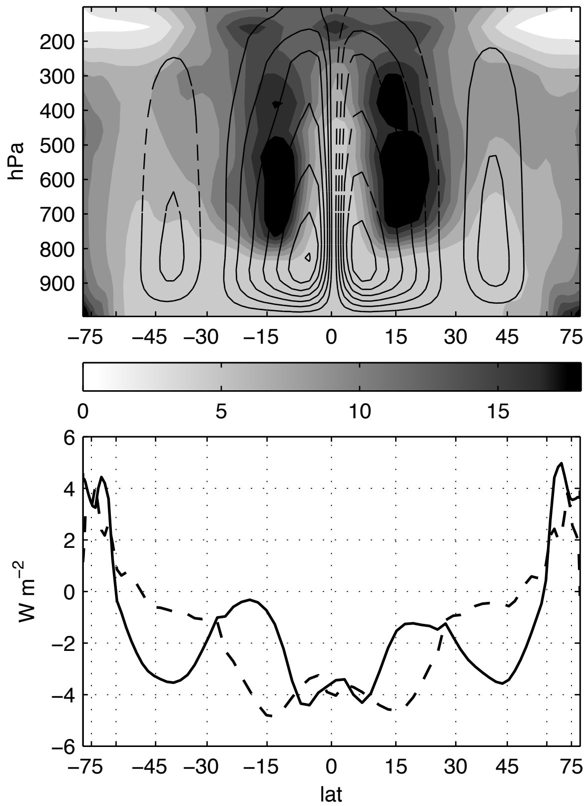

Figure 8: (a) Zonal-mean change in specific humidity (g kg^{-1}) averaged over 9 months (filled contours). Contour lines show streamlines for control climate. (b) Two estimates of the nonlinear term (W m^{-2}). Nonlinearity due to interactions amongst and within clear-sky feedbacks (solid), for the same time period as the top panel. Plotted for comparison is the residual nonlinearity (dashed) of Figure 5.

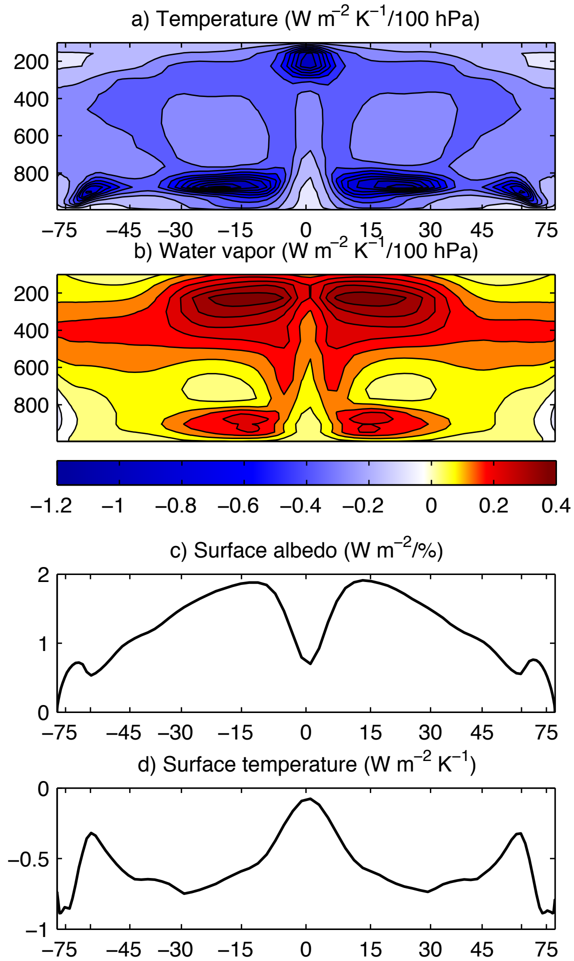

Figure B1: Zonal-mean, annual-mean radiative kernels for the GFDL aquaplanet model. (a) Temperature kernel and (b) water vapor kernel (W m^{-2} K^{-1} per 100 hPa), (c) surface albedo kernel (W m^{-2} K^{-1} per %), and (d) surface temperature kernel (W m^{-2} K^{-1}).