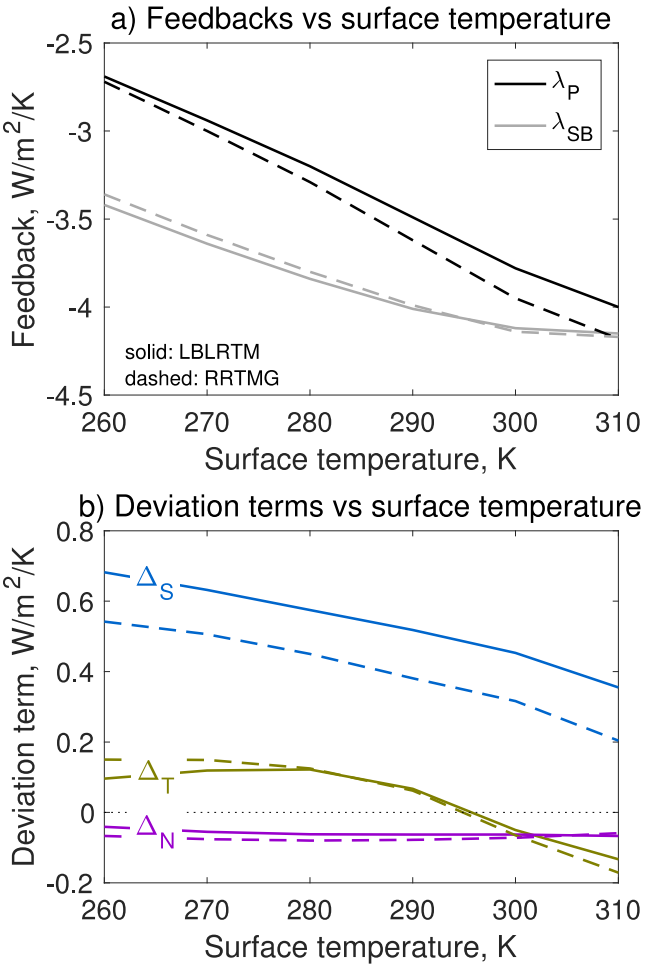
Quarterly Journal of the Royal Meteorological Society
A journal of the atmospheric sciences and physical oceanography
RMetS Royal Meteorological Society

How well do we understand the Planck feedback?

Journal:	*QJRMS*
Manuscript ID	QJ-21-0264.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Cronin, Timothy; MIT Dutta, Ishir; Massachusetts Institute of Technology, Earth, Atmospheric, and Planetary Sciences
Keywords:	Climate Change, Atmospheric Radiation, Planetary Atmospheres, Climate Feedbacks
Country Keywords:	United States

How well do we understand the Planck feedback?

Timothy W Cronin*, Ishir Dutta



A simple estimate of Earth's "no-feedback" radiative response to warming, λ_{SB}, is given by the Stefan-Boltzmann law as roughly -3.8 W m^{-2} K^{-1}. The "Planck feedback" (λ_P) widely used in analysis of climate models, however, averages -3.3 W m^{-2} K^{-1}. This 0.5 W m^{-2} K^{-1} difference has not been previously explained. In this paper we define and quantify several deviation terms that account for the difference between climate model calculations and the simple estimate based on the Stefan-Boltzmann law.

QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY

How well do we understand the Planck feedback?

Timothy W. Cronin[1] | **Ishir Dutta**[1]

[1]Program in Atmospheres, Ocean, and Climate, MIT, Cambridge, Massachusetts, USA

Correspondence
Timothy Cronin, Program in Atmospheres, Oceans, and Climate, MIT, 77 Massachusetts Ave., Cambridge, MA 02139, USA.
Email: twcronin@mit.edu

Funding information
NSF Atmospheric Chemistry grant AGS-1906719: "Advancing the Understanding of the Impacts of Wave-Induced Temperature Fluctuations on Atmospheric Chemistry."

A simple estimate of Earth's "no-feedback" radiative response to warming is given by the Stefan-Boltzmann law as $\overline{\lambda_{SB}} = -4\sigma\overline{T_e}^3$, where $\overline{T_e}$ is a global effective emission temperature. For Earth's present climate, $\overline{\lambda_{SB}} \approx -3.8$ W m^{-2} K^{-1}. The "Planck feedback" widely used in analysis of climate models, however, averages $\overline{\lambda_P} \approx -3.3$ W m^{-2} K^{-1}. This 0.5 W m^{-2} K^{-1} difference between $\overline{\lambda_P}$ and $\overline{\lambda_{SB}}$ is large compared to uncertainty in the net climate feedback, yet it has not been studied carefully. We find this difference arises mostly from the convention of calculating the Planck feedback as the top-of-atmosphere radiative flux change from vertically uniform warming of the surface and troposphere, with stratospheric temperatures unchanged (justified by differing constraints on and time scales of stratospheric adjustment). The Planck feedback is thus masked for wavelengths absorbed by the stratosphere. Other sources of difference between $\overline{\lambda_P}$ and $\overline{\lambda_{SB}}$ include covariance of the spatiotemporal warming pattern and Planck feedback, nonlinearities of averaging the derivatives of the Planck function – across wavelengths, heights, and different locations – and a non-blackbody effect of temperature-dependent gas optical properties which is included in the Planck feedback. These three additional effects each have magnitude ~0.1 W m^{-2} K^{-1}, but tend to partly cancel one another for Earth's present climate.

KEYWORDS
Climate feedbacks, Climate change, Atmospheric radiation, Planetary atmospheres

1 | INTRODUCTION

Analysis of global radiative energy balance in terms of forcings and feedbacks is a cornerstone in our understanding of global climate change. Key to such analysis is defining an appropriate "no-feedback" model for the response of global radiative balance to temperature change, to which one can add positive or negative feedbacks that amplify or dampen temperature change. The increase in blackbody radiation lost to space by a warmer atmosphere and surface provides a useful "no-feedback" model of climate change; this concept was introduced by Hansen et al. (1984) and expressed in their work as the expected temperature change for a standard radiative perturbation corresponding roughly to a doubling of CO_2. More recent studies (e.g., Bony et al., 2006; Soden and Held, 2006; Zelinka et al., 2020) denote this "no-feedback" model as the "Planck feedback," $\overline{\lambda_P}$, and express it as the change in energy gain by the climate system per degree of warming (in units of W m^{-2} K^{-1}; note that we use an overbar, $\bar{\cdot}$, to denote a global average). A natural estimate of the Planck feedback is given by the Stefan-Boltzmann law as $\overline{\lambda_{SB}} = -4\sigma\overline{T_e}^3 \approx -3.8$ W m^{-2} K^{-1} (e.g., Bony et al., 2006), where $\overline{T_e} \approx 255$ K is an effective emission temperature based on Earth's average outgoing longwave radiation (OLR ≈ 240 W m^{-2}, e.g., Loeb et al., 2018); we refer to this quantity throughout as the "Stefan-Boltzmann feedback" or "Stefan-Boltzmann estimate" (of the Planck feedback). Analysis of comprehensive general circulation models (GCMs), however, gives a Planck feedback of about -3.3 W m^{-2} K^{-1}, which is 0.5 W m^{-2} K^{-1} more positive or destabilizing than the Stefan-Boltzmann feedback (Zelinka et al., 2020; Soden and Held, 2006; Bony et al., 2006). This missing 0.5 W m^{-2} K^{-1} varies little among climate models, but has not been studied carefully and represents a notable gap in our understanding of Earth's "no-feedback" climate response. This paper seeks to close this gap and account for the deviation between the global Planck ($\overline{\lambda_P}$) and Stefan-Boltzmann ($\overline{\lambda_{SB}} = -4\sigma\overline{T_e}^3$) feedbacks.

A total deviation of 15% between $\overline{\lambda_P}$ and $\overline{\lambda_{SB}}$ might seem small, and decomposing this deviation into several components might seem like an exercise in splitting hairs – particularly since climate models agree so closely on the value of the Planck feedback. The implications would be striking, however, if the Planck feedback were 0.5 W m^{-2} K^{-1} more negative and all other feedbacks remained the same. Zelinka et al. (2020) show that the net climate feedback in climate models from the coupled model intercomparison project, phases 5 and 6 (CMIP5 and CMIP6) averages -1 W m^{-2} K^{-1}. Adding -0.5 W m^{-2} K^{-1} to this would reduce total climate sensitivity by a third, and dramatically reduce the intermodel spread in climate sensitivity as well. It seems imprudent to allow such a large unexplained gap in Earth's "no-feedback" climate sensitivity to persist without a thorough understanding of why it arises, and upon what aspects of the climate system it depends. Note that an alternative definition of the Planck feedback that was closer to the Stefan-Boltzmann estimate would not alter global climate sensitivity or its intermodel spread, but would result in attributing that $\sim$0.5 W m^{-2} K^{-1} to some combination of other feedbacks, altering our view of their relative importance to climate change.

The difference between global-mean Planck and Stefan-Boltzmann feedbacks could emerge as a result of differences between their local values, or in the process of averaging either to obtain a global number. We begin by showing that the process of spatial and temporal averaging does not explain the bulk of the discrepancy, leading us to devote most of the paper to the difference between local values of λ_P and λ_{SB}.

The global-average Planck feedback, $\overline{\lambda_P}$, is calculated as a temperature-change-weighted average of local Planck feedbacks that depend on latitude, longitude, and time (i.e., $\lambda_P = \lambda_P(\phi, \theta, t)$). Pattern covariance between local Planck feedbacks and warming will cause the global-mean Planck feedback to differ from a simple average over space and time. This effect is similar to that discussed by Back et al. (2013) with regards to fractional changes in global-mean atmospheric water vapor with global warming. Using the temperature change pattern from several climate models together with a temperature kernel, we estimate below that this deviation due to "pattern covariance" between local Planck feedback and warming likely increases $\overline{\lambda_P}$ by only $\sim$0.08 W m^{-2} K^{-1} compared to the simple spatiotemporal

average of λ_P.

One might also suspect that the nonlinearity of the Stefan-Boltzmann law could cause $\overline{\lambda_{SB}}$, based on Earth's mean outgoing longwave radiation, to differ markedly from the average across space and time of local column values of $\lambda_{SB} = -4\sigma T_e^3$. However, we find below that the difference between these two quantities due to nonlinear averaging is very small – on the order of 0.01 W m^{-2} K^{-1}.

We thus turn to an examination of why even the local Planck feedback, λ_P, calculated following current conventions for GCMs, may differ from the local Stefan-Boltzmann feedback; we identify three contributing mechanisms. The dominant cause for the 0.5 W m^{-2} K^{-1} gap between λ_P and λ_{SB} lies in a widely-used convention to calculate the Planck feedback from climate model output: λ_P is based on the increase in OLR from tropospheric and surface warming only, with no stratospheric warming. This means that the Planck feedback has almost zero contribution from spectral regions where the stratosphere is optically thick – making it considerably less negative. We refer to this dominant source of deviation of λ_P from λ_{SB} as "stratospheric masking," as the stratosphere partly masks the increase in tropopause flux from a warmer surface and troposphere.

Two smaller terms also play a role in the deviation between local Planck and Stefan-Boltzmann feedbacks, and would remain if the Planck feedback included warming of the entire column, including the stratosphere. First, the actual increase in energy loss for uniform warming of the whole column and surface is given by the nonlinear derivative of the Planck function with respect to temperature, integrated over a range of emitting temperatures (in height) and wavenumbers ($\tilde{\nu}$); this nonlinearity means that the average derivative of the flux will not equal the derivative of the average flux. This local deviation of λ_P from λ_{SB} due to nonlinear averaging echoes the issue discussed above regarding local and global values of λ_{SB}, but turns out to be somewhat more important. Second, absorption coefficients of gases – both from line and continuum absorption – depend on temperature, even with fixed concentrations. This effect causes a deviation of λ_P from λ_{SB} due to "temperature-dependent opacity," and was first discussed by Huang and Ramaswamy (2007). Using a line-by-line radiative transfer model, we find below that nonlinear averaging and temperature-dependent opacity each have magnitude ~0.1 W m^{-2} K^{-1}, but tend to cancel for Earth's present climate – although this cancellation is fortuitous and need not hold for dramatically different climate states or atmospheric compositions.

Some readers may be surprised to learn that calculations of the Planck feedback usually neglect stratospheric temperature change. This convention appears to go back at least 15 years; Bony et al. (2006) express the methodology succinctly:

> *Note that in GCM calculations, the Planck feedback parameter is usually estimated by perturbing in each grid box the tropospheric temperature at each level by the surface temperature change predicted under climate warming. Therefore this estimate does not correspond exactly to a vertically and horizontally uniform temperature change.*

This convention persists in most of the modern literature, particularly in feedback analysis that uses radiative kernels (e.g., Soden et al., 2008; Feldl and Roe, 2013; Zelinka et al., 2020). Some studies have accounted for stratospheric warming; for instance, Colman (2003) referred to a clear-sky temperature feedback of −3.9 W m^{-2} K^{-1} as being calculated from "uniform increase in temperature throughout the atmosphere (including the stratosphere)" – note that his estimated feedback parameter is larger in magnitude than those from other studies (e.g., Soden and Held, 2006; Zelinka et al., 2020). The assumption underlying the lack of stratospheric temperature change in GCM-based calculations of the Planck feedback is that the stratosphere is thermally decoupled from surface and tropospheric temperature change, and instead remains in radiative equilibrium in a perturbed climate. The new stratospheric

radiative equilibrium state, in turn, is more strongly modified by many forcing agents than it is by the temperature of the underlying troposphere and surface, and also responds more rapidly to perturbations than do the surface and troposphere. These properties of stratospheric radiative equilibrium have led to different conventions to attempt to include adjustment of stratospheric temperatures as part of the radiative forcing associated with a change in greenhouse gases (e.g., Hansen et al., 1997). A more consistent choice for the Planck feedback would be to compute the OLR change due to a warmer troposphere and surface, while keeping the overlying stratosphere in radiative equilibrium with the altered upward longwave flux at the tropopause. This choice would pose its own challenges, however, as it requires numerous offline radiative transfer calculations and could not be done easily with a standard radiative kernel.

In this paper, we define and quantify the sources of the ~0.5 W m^{-2} K^{-1} deviation between Planck and Stefan-Boltzmann feedbacks. Throughout, we use a sign convention that negative feedbacks or deviation terms are stabilizing, whereas positive feedbacks or deviation terms are destabilizing. We begin by using climate model output and a widely-used radiative kernel from a climate model to examine how two artifacts of global averaging might cause the global-mean Planck and Stefan-Boltzmann feedbacks to deviate from simple spatiotemporal means of their local values (Section 2). We then turn to the local deviation between λ_P and λ_{SB}, beginning with a description of how different deviation terms can be isolated from one another, and then describing numerical calculations with a radiative transfer model to perform this task (Section 3). Calculations for an example atmospheric sounding reveal the dominant role of stratospheric masking in making λ_P less negative than λ_{SB}, and the nontrivial but generally cancelling roles of nonlinear averaging and temperature-dependent opacity (Section 4.1). We explore how the deviation between λ_P and λ_{SB} depends on the surface and atmospheric temperatures, which shows that over a third of the difference in local Planck feedbacks between low and high latitudes on Earth might be caused by a systematically smaller difference between λ_P and λ_{SB} for warmer surfaces, rather than solely by the variation of T_e with surface temperature (Section 4.2). We present calculations for an atmosphere with a dramatically different composition of greenhouse gases, which leads to different relative balances of the deviation terms (Section 4.3), and conclude with a discussion of limitations of this work and future directions (Section 5).

2 | GLOBAL AVERAGING

The global-mean Planck feedback, $\overline{\lambda_P}$, is a weighted average of local feedbacks based on local temperature changes (in latitude, longitude, and time), so any covariance between the warming pattern and local Planck feedback will cause $\overline{\lambda_P}$ to differ from a simple (area-weighted) global average over space and time. Furthermore, the global-mean Stefan-Boltzmann feedback $\overline{\lambda_{SB}} = -4\sigma\overline{T_e}^3$ – where $\overline{T_e}$ is an effective emission temperature for the planet defined from global-mean OLR – will not generally match the global average of local Stefan-Boltzmann feedbacks. Thus, one might suppose that local values of λ_P and λ_{SB} could match closely, and that the bulk of the discrepancy between global-mean values of the two quantities occurs due to these two artifacts of global averaging. In this section we show that only a ~0.1 W m^{-2} K^{-1} discrepancy can be explained by global averaging, leaving the bulk of the difference between $\overline{\lambda_P}$ and $\overline{\lambda_{SB}}$ to be explained by local deviation between λ_P and λ_{SB}.

2.1 | Pattern covariance

Most climate models show warming patterns that maximize near the north pole in winter, where the Planck feedback is also anomalously positive, so we expect the temperature-change-weighted Planck feedback to be less negative than its simple area-averaged value. Quantifying this difference requires examining climate model output, together with

information about how the Planck feedback varies in space and time.

We use the Soden et al. (2008) all-sky radiative kernel from the GFDL model, which has a spatial resolution of 2 degrees latitude by 2.5 degrees longitude, 17 vertical levels, and monthly temporal resolution. In order to calculate the local Planck feedback, we sum the temperature kernel from levels between the surface and tropopause, with tropopause pressure defined by the simple approximate expression $p_{tpp} = 300 - 200\cos\phi$ hPa. We denote the temperature-change-weighted global-mean Planck feedback as $\overline{\lambda_P}^{(1,1,1)}$, with the superscripts following the overline indicating that covariance between warming pattern and local Planck feedback is included for latitude, longitude, and time of year. Similarly, we denote the simple area-weighted mean Planck feedback as $\overline{\lambda_P}^{(0,0,0)}$, to indicate that it includes no covariance between warming and local feedback. No warming pattern is required to calculate $\overline{\lambda_P}^{(0,0,0)}$, as it represents the simple mean of local Planck feedbacks. Using the Soden et al. (2008) kernel, and with the tropopause defined as above, we find $\overline{\lambda_P}^{(0,0,0)} = -3.263\ \mathrm{W\,m^{-2}\,K^{-1}}$.

We use temperature change output from the abrupt CO_2 quadrupling and historical scenarios of the Climate Model Intercomparison Project, phase 6, taking the difference between the last 10 years of the abrupt $4\times CO_2$ simulation and the first 10 years of the historical simulation to represent the temperature change pattern. We use the mean and inter-model spread between 19 models as representative of a predicted climate change warming pattern; the models used are listed in Table A1. Monthly temperature change patterns are interpolated to the same horizontal grid as the radiative kernel to compute the covariance between the warming pattern and local Planck feedback. Averaging across these simulations, we find $\overline{\lambda_P}^{(1,1,1)} = -3.185\ \mathrm{W\,m^{-2}\,K^{-1}}$ – about $0.08\ \mathrm{W\,m^{-2}\,K^{-1}}$ more positive than the simple mean $\overline{\lambda_P}^{(0,0,0)}$. This leads us to the primary result of this section: the deviation of the global Planck feedback from a simple average due to covariance between the warming pattern and local feedback is a non-negligible $0.08\ \mathrm{W\,m^{-2}\,K^{-1}}$ compared to the total $0.5\ \mathrm{W\,m^{-2}\,K^{-1}}$ gap between $\overline{\lambda_P}$ and $\overline{\lambda_{SB}}$, but it cannot explain the bulk of the discrepancy. In Appendix A (and Table A1), we show that this term occurs mostly due to the covariance of Planck feedback and warming in latitude, with secondary effects from the combined covariance of the two quantities in latitude and time of year, and negligible effects from including covariance in longitude.

2.2 | Global averaging of the Stefan-Boltzmann feedback

The global-mean effective emission temperature is defined based on the global-mean OLR, $\overline{F_0}$, as $\overline{T_e} = (\overline{F_0}/\sigma)^{1/4}$. This implies that $\overline{\lambda_{SB}} = -4\sigma^{1/4}\overline{F_0}^{3/4}$. Local values of T_e, on the other hand, are defined using local outgoing fluxes F_0, so the average of local Stefan-Boltzmann feedbacks is given by $-4\sigma^{1/4}\overline{F_0^{3/4}}$. Since the 3/4 power is concave down, the Stefan-Boltzmann feedback based on global-mean OLR will always be more negative than the global-mean of local Stefan-Boltzmann feedbacks. Taking $F_0 = \overline{F_0} + F_0'$, and expanding $F_0^{3/4}$ as:

$$F_0^{3/4} = \left(\overline{F_0} + F_0'\right)^{3/4} \approx \overline{F_0}^{3/4}\left(1 + \frac{3F_0'}{4\overline{F_0}} - \frac{3F_0'^2}{32\overline{F_0}^2} + \ldots\right), \tag{1}$$

which gives a difference:

$$-4\sigma^{1/4}\left(\overline{F_0}^{3/4} - \overline{F_0^{3/4}}\right) \approx -4\sigma^{1/4}\overline{F_0}^{3/4}\left(\frac{3\overline{F_0'^2}}{32\overline{F_0}^2}\right). \tag{2}$$

In words, the relative error incurred by taking the global-average of the OLR first, rather than calculating local Stefan-Boltzmann feedbacks and then averaging them, is given roughly by 3/32 times the spatiotemporal variance of OLR,

divided by the square of the global-mean OLR. Using daily OLR observations from satellite (uninterpolated OLR data provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website at `https://psl.noaa.gov/data/gridded/data.uninterp_OLR.html`), we calculate $\overline{F_0} = 232.1$ W m^{-2}, and $\overline{F_0'^2} = 1659$ W^2 m^{-4}, which gives an estimated error of 0.01 W m^{-2} incurred by globally averaging OLR before calculating the Stefan-Boltzmann feedback. The binomial expansion used here also turns out to be an extremely good approximation, with values of $-4\sigma\overline{F_0^{3/4}}$ and $-4\sigma\overline{F_0}^{3/4}(1 - \frac{3}{32}\overline{F_0'^2}/\overline{F_0}^2)$ differing by less than 0.001 W m^{-2}. To recap: we consider this ~0.01 W m^{-2} artifact of global averaging of the Stefan-Boltzmann feedback to be negligibly important and do not discuss it further – for Earth, its value is so small because the 3/4 power is a weakly nonlinear function, and because the relative variation in OLR is small compared to its mean value.

3 | LOCAL DEVIATION TERMS

3.1 | Definitions

The local difference between the Planck and Stefan-Boltzmann feedbacks can be attributed to three sources: stratospheric masking, temperature-dependent opacity, and nonlinear averaging. Following GCM conventions, the Planck feedback λ_P is an integral over wavenumber $\tilde{\nu}$ of the top-of-atmosphere flux change $\delta F_0^{\tilde{\nu}}$ (where the subscript $_0$ indicates the top-of-atmosphere, and the superscript $^{\tilde{\nu}}$ indicates that the flux depends on wavenumber) per unit vertically uniform warming of the surface and troposphere, denoted δT_T. These flux changes can be further decomposed into contributions from the increase in Planck source function with warming $(\delta F_0^{\tilde{\nu}})_{\text{Planck}}$, and contributions from changes in gas optical properties (which may be either sign), $(\delta F_0^{\tilde{\nu}})_{\text{optics}}$:

$$\lambda_P = -\int_0^\infty \left[\left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_T}\right)_{\text{Planck}} + \left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_T}\right)_{\text{optics}}\right] d\tilde{\nu}. \tag{3}$$

Adding and subtracting λ_{SB} from the right-hand side (and writing its spectrally-resolved form as $-\pi dB^{\tilde{\nu}}(T_e)/dT$), adding and subtracting the flux derivative per unit stratospheric temperature change due to the stratospheric Planck source, $(\delta F_0^{\tilde{\nu}}/\delta T_S)_{\text{Planck}}$ (where δT_S denotes only stratospheric warming), and regrouping terms, gives:

$$\lambda_P = \lambda_{SB} \underbrace{- \int_0^\infty \left[\left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_T}\right)_{\text{Planck}} + \left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_S}\right)_{\text{Planck}} - \pi \frac{dB^{\tilde{\nu}}(T_e)}{dT}\right] d\tilde{\nu}}_{\Delta_N} \underbrace{+ \int_0^\infty \left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_S}\right)_{\text{Planck}} d\tilde{\nu}}_{\Delta_S} \underbrace{- \int_0^\infty \left(\frac{\delta F_0^{\tilde{\nu}}}{\delta T_T}\right)_{\text{optics}} d\tilde{\nu}}_{\Delta_T}. \tag{4}$$

In equation 4, each underbrace defines a deviation term:

$$\begin{aligned} \lambda_P &= \lambda_{SB} + \Delta_N + \Delta_S + \Delta_T \\ \Delta_N &: \text{nonlinear averaging} \\ \Delta_S &: \text{stratospheric masking} \\ \Delta_T &: \text{temperature – dependent opacity.} \end{aligned} \tag{5}$$

The temperature-dependent opacity term is not included in the Stefan-Boltzmann feedback and represents physics of gas absorption coefficients that strengthen or weaken with warming – distinct from the variation of blackbody radiation with temperature. The stratospheric masking deviation is the derivative of the top-of-atmosphere flux with

respect to stratospheric temperature due to the changes in the stratosphere's Planck source function – this term is "missing" from the standard definition of λ_P because stratospheric temperatures are held constant. The nonlinear averaging deviation is the difference between the whole-column changes in Planck source function with warming and the derivative of blackbody emission with respect to temperature, evaluated at T_e. The nonlinear averaging deviation includes contributions both from nonlinear averaging over heights at a given wavenumber and nonlinear averaging across wavenumbers; in Appendix B we show that the former term is usually small.

3.2 | Radiative transfer calculations

We use calculations with the line-by-line radiative transfer model LBLRTM (Clough et al., 2005) to quantify λ_P, λ_{SB}, and the three local deviation terms. We use a spectral resolution of $\delta\tilde{\nu} \sim 0.01$ cm^{-1} over the thermal infrared from $\tilde{\nu} = 10 - 3500$ cm^{-1}, so a few hundred thousand monochromatic radiative transfer calculations are done for each profile. This allows for the effects of individual lines (typically with widths ~0.1 cm^{-1} at sea-level pressure) to be resolved. Output is also averaged over 5 cm^{-1} intervals for purposes of plotting.

We use a reference atmospheric profile with a surface temperature of 290 K, a moist pseudoadiabatic troposphere with 80% relative humidity, and a stratosphere with a constant lapse rate $dT/dz = 1.8$ K km^{-1} above the height where temperatures fall to a specified tropopause temperature of 200 K (for this profile, 12.5 km). The surface pressure is set to 1000 hPa, and the 1000-hPa air temperature equals the surface temperature. The specific humidity in the stratosphere is set to a uniform 5 ppmv (similar to observed mid-stratospheric values; e.g., Oman et al., 2008), and the ozone profile follows the gamma distribution in pressure from Wing et al. (2018):

$$q_{O_3} = 3.6478 \times 10^{-6} p^{0.83209} e^{-p/11.3515}, \tag{6}$$

where q_{O_3} is the ozone volume mixing ratio, and p is the atmospheric pressure in hPa. The reference atmospheric profile includes 400 ppmv of CO_2, and no other well-mixed greenhouse gases. These choices are intended to represent a compromise between a profile reasonably close to global-average conditions, while also being simple and easily generalized to different surface temperatures. The dry atmosphere is taken to be 79% N_2 and 21% O_2 (relevant for pressure-broadening of lines). The vertical grid spacing is 500 m, fluxes are integrated to a height of 50 km, and gas absorber amounts are scaled by a factor of 5/3 to account for the slant path taken by thermal radiation through the atmosphere. We perform clear-sky calculations only, but will discuss later the potential impact of clouds.

We use several calculations with the standard LBLRTM code, as well as with a modified code where the temperatures seen by gas optics only are increased by 1 K at all included levels, to isolate the contribution from each term:

1. A top-of-atmosphere flux control calculation gives a reference infrared flux spectrum $F_0^{\tilde{\nu}}$ and thus defines $T_e = (\int_0^\infty F_0^{\tilde{\nu}} d\tilde{\nu}/\sigma)^{1/4}$ and $\lambda_{SB} = -4\sigma T_e^3$.
2. A tropopause-flux control calculation defines an infrared flux spectrum at the tropopause, $F_T^{\tilde{\nu}}$ (the subscript $_T$ indicating tropopause flux).
3. A top-of-atmosphere flux, troposphere-cooled calculation uses tropospheric and surface temperatures 1 K cooler than the control calculation. The flux difference between this calculation and the control gives $\lambda_P = \delta F_0^{\tilde{\nu}}/\delta T_T$ following Equation 3.
4. A tropopause-flux, troposphere-cooled calculation gives an estimate of $\delta F_T^{\tilde{\nu}}/\delta T_T$, from which we calculate the

(spectrally-resolved) stratospheric emissivity as:

$$\epsilon_S = 1 - \frac{\delta F_0^{\tilde{\nu}}/\delta T_T}{\delta F_T^{\tilde{\nu}}/\delta T_T}. \tag{7}$$

5. A tropopause-flux, troposphere-cooled calculation using the modified code to calculate gas optical properties from the control temperature profile. Comparing this calculation with the tropopause-flux troposphere-cooled calculation (where gas optical properties also see the temperature perturbation) then allows us to calculate $(\delta F_T^{\tilde{\nu}}/\delta T_T)_{\text{optics}}$ directly, and $(\delta F_T^{\tilde{\nu}}/\delta T_T)_{\text{Planck}}$ as the difference $\delta F_T^{\tilde{\nu}}/\delta T_T - (\delta F_T^{\tilde{\nu}}/\delta T_T)_{\text{optics}}$. The top-of-atmosphere flux differences $(\delta F_0^{\tilde{\nu}}/\delta T_T)_{\text{Planck}}$ and $(\delta F_0^{\tilde{\nu}}/\delta T_T)_{\text{optics}}$ can then be calculated by multiplying the respective tropopause flux differences by $(1 - \epsilon_S)$ to account for stratospheric attenuation.
6. A top-of-atmosphere flux, full-column cooled calculation using the modified code to calculate gas optical properties from the control temperature profile. Comparing this calculation with the control case gives $(\delta F_0^{\tilde{\nu}}/\delta T_T)_{\text{Planck}} + (\delta F_0^{\tilde{\nu}}/\delta T_S)_{\text{Planck}}$, and thus allows us to isolate the change in stratospheric Planck source term by subtracting our previous calculation of $(\delta F_0^{\tilde{\nu}}/\delta T_T)_{\text{Planck}}$.

4 | RESULTS FOR LOCAL DEVIATION TERMS

4.1 | Reference atmosphere

The outgoing longwave radiation shows absorption lines across the entire spectrum, but most strongly from a CO_2 rotational-vibrational band between 550 and 750 cm^{-1}, from an O_3 rotational-vibrational band near 1050 cm^{-1}, from the pure rotational band of H_2O at 0-500 cm^{-1} and a rotational-vibrational band of H_2O from 1250-2000 cm^{-1} (Figure 1a). This reference atmosphere has OLR=261.3 W m^{-2} (~5-10 W m^{-2} smaller than the observed clear-sky global-mean, e.g., Loeb et al., 2018), an effective emission temperature $T_e = 260.6$ K, and thus a local Stefan-Boltzmann feedback of -4.01 W m^{-2} K^{-1}.

A calculation with the surface and troposphere warmed by 1K, however, shows that the Planck feedback is less negative than λ_{SB} by 0.52 W m^{-2} K^{-1}: $\lambda_P = -3.49$ W m^{-2} K^{-1} (Figure 1b, red). The Planck feedback is nearly zero in parts of the spectrum where the stratosphere is optically thick, most notably in the CO_2 and O_3 bands, but also to a lesser extent in the stronger parts of H_2O bands. The stratospheric masking deviation, $\Delta_S(\tilde{\nu}) = (\delta F_0^{\tilde{\nu}}/\delta T_S)_{\text{Planck}}$, acts as a positive feedback in these spectral regions, and accounts almost entirely for the spectrally-integrated difference of 0.52 W m^{-2} K^{-1} between λ_P and λ_{SB} (Figure 1b, blue). The temperature-dependent opacity deviation $\Delta_T(\tilde{\nu}) = -(\delta F_0^{\tilde{\nu}}/\delta T_T)_{\text{optics}}$, integrates across wavenumbers to a small positive feedback of 0.07 W m^{-2} K^{-1}, with positive contributions on the flanks of the CO_2 and water vapor bands outweighing negative contributions within the atmospheric window region (particularly from 800-1000 cm^{-1}), where continuum absorption by H_2O dominates over line absorption (Figure 1b, gold). The edges of CO_2 and H_2O bands have line strengths that depend particularly strongly on temperature because their lower-level states have high rotational quantum numbers and thus high lower-level energies, so molecular populations in the absorbing lower-level quantum state increase rapidly with warming. Continuum absorption, on the other hand, decreases with temperature if gas partial pressures are held fixed with warming. Note as well that the spectral features of our temperature-dependent opacity deviation term compare closely to the "absorptivity effect" first discussed by Huang and Ramaswamy (2007) (see their Figure 4). We find a smaller magnitude of Δ_T, possibly because Huang and Ramaswamy (2007) also include stratospheric warming. The nonlinear averaging deviation, $\Delta_N(\tilde{\nu})$, given by the difference between purple and black lines in Figure 1b, integrates to a small negative feedback of -0.06 W m^{-2} K^{-1}

and thus almost cancels Δ_T.

Cumulative integrals of each deviation in wavenumber (Figure 1c) make it easier to visually determine what regions of the spectrum contribute most to each term. The stratospheric masking term derives about ~60% of its value from the CO_2 band, and about ~20% each from H_2O, and O_3 bands (Figure 1c, blue). The temperature-dependent opacity term, Δ_T, derives about half of its value from CO_2 bands and half from H_2O bands (Figure 1c, gold). Finally, the nonlinear averaging term, Δ_N (given by the difference between purple and black lines in Figure 1b), has a large magnitude across the spectrum, since most wavenumbers do not have brightness temperatures ($T_b^{\tilde{\nu}}$) close to T_e. The cumulative integral of Δ_N shows that negative contributions in the atmospheric window region from 800-1300 cm^{-1}, where $T_b^{\tilde{\nu}u} > T_e$, are only partly offset by positive contributions in strongly absorbing bands of CO_2, H_2O, and O_3, where $T_b^{\tilde{\nu}u} < T_e$ (Figure 1c, purple). Notably, the positive contributions to Δ_N by strongly absorbing bands are not proportional to the greenhouse effect of each band, or the amount by which it reduces OLR. For example, the combination of the H_2O rotational and CO_2 rotational-vibrational bands contribute about 0.2 W m^{-2} K^{-1} to Δ_N, which is similar to the impact of the H_2O rotational-vibrational band, but the two former bands have a much larger greenhouse effect than the latter one. This mismatch arises because the high-wavenumber tail of the Planck function accounts for a much larger share of $dB^{\tilde{\nu}}(T)/dT$ as compared to $B^{\tilde{\nu}}(T)$, so absorbers at high wavenumbers matter more in a relative sense for the Planck feedback than they do for OLR.

4.1.1 | Calculations with RRTMG

Line-by-line calculations show how each deviation term arises from specific absorption features, and avoids complications from band-averaging of fluxes before or after evaluating flux derivatives. For computational efficiency, climate models typically use radiation parameterizations that approximately solve atmospheric radiative transfer equations at many fewer wavenumbers. For example, RRTMG (the Rapid Radiative Transfer Model for GCMs, Clough et al., 2005) – a widely used scheme in climate models – uses the correlated-k approximation, in which the thermal spectrum is first broken up into broad bands (16 bands for RRTMG from 10-3250 cm^{-1}; for example, one strong CO_2 band spans 630-700 cm^{-1}), and a small number of full radiative transfer calculations are then performed by grouping wavenumbers in each band with similar gas absorption coefficients (called g-points). Thus, in RRTMG, only 140 radiative transfer calculations are done for each profile, and the temperature-dependence of absorption coefficients is implemented by lookup tables rather than by explicit calculations of line strengths. The benchmark accuracy of RRTMG relative to LBLRTM is ~1 W m^{-2} for net longwave fluxes at any altitude (Clough et al., 2005). To test whether the approximations in RRTMG matter for the deviation terms, we have repeated the above calculations but using RRTMG rather than LBLRTM.

For the reference atmospheric profile, agreement between the two models is good but not perfect. The OLR and Stefan-Boltzmann feedback in RRTMG compare quite closely (within 1%) with the calculations from LBLRTM, and each deviation term in RRTMG also matches the sign and relative magnitude of that found in LBLRTM (Figure 2). The Planck feedback from RRTMG, however, is more negative than that calculated with LBLRTM by 0.13 W m^{-2} K^{-1}, with Δ_S and Δ_N both being slightly lower in RRTMG than in LBLRTM. The discrepancy is small for the nonlinear averaging term, and the use of band-averaged data in RRTMG limits our ability to explore it further. The stratospheric masking term seems smaller in RRTMG than in LBLRTM because the emissivity of the stratosphere in the O_3 band and the wings of the CO_2 band is underestimated by RRTMG. Such underestimation of stratospheric opacity, particularly in the O_3 band, is a documented error of the correlated-k approximation, likely associated with "blurring," whereby the sorting of wavenumbers by absorption coefficient is imperfectly correlated at different heights in the atmosphere (Fu and Liou, 1992).

Overall, the results of these calculations suggest that the radiative transfer code used in many GCMs is sufficient to capture the physics and general size of the deviation terms discussed above, but that stratospheric opacity may be underestimated in GCMs that use RRTMG or other radiative transfer schemes with the correlated-k approximation.

4.2 | Temperature-dependence of the deviations

Is the latitudinal variation of λ_P mostly captured by variations in T_e, or do the deviation terms also vary systematically with surface temperature? At high surface temperatures, we expect the stratosphere to be physically thinner and the value of the stratospheric masking term smaller, but it is less clear how the other two deviation terms might vary with surface warming. To explore these questions, we use calculations with atmospheres over surface temperatures T_g ranging from 260-310 K, using both LBLRTM and RRTMG. Temperature profiles again follow a moist adiabat in the troposphere up to a tropopause temperature of 200 K, and temperatures increase with height at 1.8 K km^{-1} in the stratosphere. Tropospheric relative humidity, stratospheric specific humidity, CO_2, and O_3 profiles are all identical to the reference calculation.

The Planck feedback varies more with surface temperature than does the Stefan-Boltzmann feedback, with λ_P becoming closer to λ_{SB} at high surface temperatures (Figure 3a). This greater sensitivity to temperature indicates that the sum of the deviation terms $\Delta_S + \Delta_T + \Delta_N$ becomes less positive with surface warming, and arises mainly from stratospheric masking and temperature-dependent opacity terms; nonlinear averaging varies comparatively little with surface temperature (Figure 3b). As the surface warms at constant tropopause temperature, the stratosphere thins, stratospheric emissivity decreases, and Δ_S decreases. Across all surface temperatures Δ_S is smaller for RRTMG than for LBLRTM, but the sensitivity to surface temperature is similar in both models. The temperature-dependence of Δ_T occurs due to the competition between the H_2O continuum, which decreases in optical thickness with warming, and other absorption bands, which mostly increase in optical thickness with warming. Continuum absorption becomes important only for $T_g \geq 290$K, switching the sign of Δ_T from positive to negative: specifically, the continuum region from 800-1300 cm^{-1} contributes about 0.01 W m^{-2} K^{-1} to Δ_T for surface temperatures of 260-280 K, -0.03 W m^{-2} K^{-1} for $T_g = 290$ K, -0.13 W m^{-2} K^{-1} for $T_g = 300$ K, and -0.2 W m^{-2} K^{-1} for $T_g = 310$ K.

Using the $T_g = 260$ K and $T_g = 300$ K calculations from LBLRTM as representative of the pole-equator temperature difference on Earth, the difference between λ_{SB} values at these two surface temperatures is 0.7 W m^{-2} K^{-1}, whereas the difference in λ_P is 1.09 W m^{-2} K^{-1}. These results indicate that the deviation terms increase $|d\lambda_P/dT_g|$ by over 50% as compared to $|d\lambda_{SB}/dT_g|$. In LBLRTM, stratospheric masking accounts for about 60% of this greater sensitivity of the Planck feedback to surface temperature, temperature-dependent opacity accounts for 35%, and nonlinear averaging the remaining 5%. Calculations with RRTMG also show that $|d\lambda_P/dT_g|$ exceeds $|d\lambda_{SB}/dT_g|$, with the parameterized temperature-dependent opacity accounting for a larger fraction of the effect in that model. To our knowledge, this dependence of the Planck feedback on surface temperature, which is systematic and unrelated to T_e, is a previously unrecognized facet of climate feedback analysis.

4.3 | Dramatically different atmospheric states

As an example of how a different atmospheric composition could alter the Planck feedback, we have done a calculation with all H_2O removed from the reference case, but with increased concentrations of CO_2 and a modified temperature structure tuned to give the same OLR=261.3 W m^{-2} as the reference case. The modified atmosphere has 19.5% CO_2, no O_3, a dry-adiabatic lapse rate in the troposphere, and an isothermal stratosphere at a temperature of 200 K. These choices give a markedly different OLR spectrum: at wavenumbers of 200-500 cm^{-1}, emission looks nearly like that of a

blackbody, but at higher wavenumbers, several CO_2 bands (from about 550-800, near 950, 1100, 1250, and 1350 cm^{-1}) cut deeply into the outgoing flux (Figure 4a). By construction, this example has the same value of $\lambda_{SB} = -4.01$ W m^{-2} K^{-1}, but now $\lambda_P = -2.99$ W m^{-2} K^{-1}, a further 0.5 W m^{-2} K^{-1} less negative than the reference case. Stratospheric masking is still the largest deviation term, with a value of 0.54 W m^{-2} K^{-1} (slightly larger than the reference case). Temperature-dependent opacity and nonlinear averaging deviations, however, differ much more relative to our calculations above: $\Delta_T = 0.39$ W m^{-2} K^{-1} is larger by nearly a factor of six, and $\Delta_N = 0.09$ W m^{-2} K^{-1} has changed signs and increased slightly in magnitude. The stratospheric masking term is still large because the increased abundance of CO_2 makes the stratosphere optically thick across more of the spectrum, which compensates for a smaller Planck source from a cooler stratosphere and lack of stratospheric opacity in H_2O bands. The temperature-dependent opacity term is large because the flanks of the CO_2 bands (near 550, 800, 950, and 1100 cm^{-1} in Figure 4) have line strengths that are particularly sensitive to temperature, and because there is no cancelling contribution from weakening H_2O continuum absorption with warming. Finally, the nonlinear averaging term is positive because the primary "atmospheric window" where $T_b^{\tilde{\nu}} > T_e$ now lies at low wavenumbers where the value of $dB^{\tilde{\nu}}/dT$ is smallest relative to $B^{\tilde{\nu}}$, and higher wavenumbers are more evenly split between brightness temperatures above and below T_e.

This example shows that difference between the Planck and Stefan-Boltzmann feedbacks could be much larger, and that the cancellation between temperature-dependent opacity and nonlinear averaging deviations found in the present climate need not hold for different atmospheric compositions. The deviation terms Δ_T and Δ_N in this example could be viewed as important positive feedbacks of their own, with their sum being comparable in magnitude to the surface albedo or cloud feedbacks in current climate models (e.g. Zelinka et al., 2020). This specific example also suggests that worlds with dry, CO_2-rich atmospheres could show considerably greater climate sensitivity than we would estimate from the Stefan-Boltzmann feedback.

5 | DISCUSSION AND CONCLUSIONS

Using climate model patterns for warming and the Planck feedback from a radiative kernel, we have shown that the two possible artifacts of global averaging do not suffice to explain the ~0.5 W m^{-2} K^{-1} difference between global-mean Planck and Stefan-Boltzmann feedbacks. We find that the covariance of warming pattern and local Planck feedbacks causes the global-mean Planck feedback, $\overline{\lambda_P}$, to be about 0.1 W m^{-2} K^{-1} more positive than the simple areal average of local Planck feedbacks, and that nonlinearity of global averaging is insignificant for calculating the global-mean Stefan-Boltzmann feedback.

Following on these results, we have used single-column calculations with both a line-by-line and a correlated-k radiative transfer model to show that the deviation between local Planck and Stefan-Boltzmann feedbacks can be accounted for by three deviation terms – related to stratospheric masking, temperature-dependent opacity, and nonlinear averaging. We find that stratospheric masking dominates, increasing λ_P relative to λ_{SB} by ~0.5 W m^{-2} K^{-1} near global-mean surface temperatures, with a smaller contribution at higher surface temperatures when the stratosphere is thinner and a larger contribution at lower surface temperatures when the stratosphere is thicker. Temperature-dependent opacity and nonlinear averaging each have magnitudes of about 0.1 W m^{-2} K^{-1} but opposing signs; temperature-dependent opacity acts as a positive feedback whereas nonlinear averaging acts as a negative feedback near present global-mean surface temperatures. The stratospheric masking and temperature-dependent opacity terms both depend on surface temperature, making λ_P more sensitive to local surface temperatures than λ_{SB} is, and potentially accounting for over a third of the meridional variation in the Planck feedback.

For Earth-like conditions, stratospheric masking is the largest deviation term between the local Planck and Stefan-

Boltzmann feedbacks. Stratospheric masking can be included in a simple view of the total clear-sky longwave feedback – which includes Planck, water vapor, and lapse rate components – by slightly reframing the central result of Ingram (2010). Ingram (2010), highlighting the seminal work of Simpson (1928), clarified that to first order, spectral regions where water vapor makes the atmosphere optically thick show almost zero change in outgoing infrared flux with warming (at constant relative humidity), while all other wavenumbers will show a 'Planckian' increase in flux with warming (following $dB^{\tilde{\nu}}(T_b^{\tilde{\nu}})/dT$). More recently, Jeevanjee et al. (2021) explicitly showed that this effect is naturally captured when computing climate feedbacks with relative humidity held fixed. Accounting for stratospheric masking slightly revises this rule: any spectral regions that are not optically thick either to water vapor *or in the stratosphere* will to first-order show a 'Planckian' increase in flux with warming. Stratospheric optical thickness thus emerges as an important aspect of the clear-sky net longwave feedback, which to our knowledge has not been previously considered.

There are at least two other possible conventions for the Planck feedback that eliminate the stratospheric masking deviation term and thus lead to a much smaller difference of the Planck feedback from the Stefan-Boltzmann estimate. First, one could easily compute the Planck feedback from uniformly warming the whole column, including the stratosphere. Second, one could compute the tropopause flux changes from surface and tropospheric warming. We have done calculations for both choices, and found that Δ_N and Δ_T are each slightly larger in magnitude, but still largely cancel, leaving either of these alternative Planck feedbacks within ~0.05 W m^{-2} K^{-1} of an appropriately defined Stefan-Boltzmann feedback. Both conventions are still confounded, however, by stratospheric adjustment. In the case of full column warming, the lack of stratospheric warming (or stratospheric cooling) associated with climate change would appear as a positive component of the lapse-rate feedback rather than as a positive component of the Planck feedback. This choice might seem appealing as it captures the temperature change structure of the full column, but it is also potentially problematic because it could (if done without care) classify the adjustment of stratospheric temperatures to forcing agents as a feedback, even though it occurs much faster than surface temperature change. In the case of tropopause fluxes, a rigorous treatment would require also considering the effect on tropopause fluxes of stratospheric temperature adjustment to a warmer troposphere – a complement to the negative top-of-atmosphere feedback considered by Wang and Huang (2020) that would instead appear as a positive "stratospheric temperature feedback" at the tropopause. If the stratosphere behaved as an optically thin gray absorber, then one would expect the stratosphere to warm in response to an increase in absorbed infrared flux from below such that it radiated half of that perturbation flux upward and half downward – suggesting that this "stratospheric temperature feedback" might be about half the size of the stratospheric masking deviation Δ_S. However, the stratosphere is far from a gray absorber; the sensitivity of stratospheric temperature and top-of-atmosphere fluxes to tropopause flux changes is nuanced and depends on both the spectral distribution of stratospheric emissivity and the base-state stratospheric temperature profile. Neither choice of an alternative convention for calculating the Planck feedback provides a clear advantage, though both could clarify the role of the stratosphere in climate feedbacks more explicitly, similar to the intended goal of our Δ_S.

This paper has focused on clear-sky conditions; cloud cover would modify the magnitudes of some of the local deviation terms. The stratospheric masking term would be unchanged by clouds, since the stratospheric emission missing from the Planck feedback under the assumption of no stratospheric temperature change is entirely independent of (tropospheric) clouds. By reducing the outgoing flux and smoothing its spectrum towards that of a blackbody at cloud-top temperature, optically thick cloud layers – especially in the upper troposphere – would likely reduce the magnitudes of both Δ_T and Δ_N. Low clouds are unlikely to greatly alter the picture presented in this paper, though if they were to fall above a moist boundary layer with strong H_2O continuum absorption in the Tropics, they might make the temperature-dependent opacity term more positive at high temperatures. Clouds could in some situations make Δ_N more positive by modifying the part of the nonlinear averaging term arising from averaging over heights, as discussed in

Appendix B. This would be particularly likely for cold clouds of optical thickness ~1, or in the case of a scene with partial cover by high thick clouds.

A further uncertainty and possibility for inter-model spread in the Planck feedback arises from the water vapor continuum. A decrease in water vapor continuum absorption with temperature (holding vapor pressure fixed) is empirically well-established, and incorporated into the prevailing MT_CKD continuum model which is used by both LBLRTM and RRTMG (Clough et al., 2005), as well as by many other radiation codes in climate models (e.g., Oreopoulos et al., 2012). Nevertheless, experimental studies disagree on how strongly continuum absorption decreases with warming (Cormier et al., 2005), and the physical mechanisms underlying the mid-infrared water vapor continuum – whether related primarily to water vapor dimers or to far wings of strong lines in rotational and rotational-vibrational bands – still remain controversial (e.g., Shine et al., 2012). Uncertainty in water vapor continuum absorption could impact the temperature-dependent opacity deviation term, especially at high temperatures where the continuum region dominates Δ_T.

Do these findings matter for climate sensitivity and the sum of climate feedbacks? In a narrow sense, the answer is no; we have merely explained aspects of the Planck feedback that, although poorly understood, are incorporated into climate model analysis and show relatively little inter-model spread. In a broader sense, however, this work clarifies how several properties of the climate system – stratospheric optical depth, temperature-dependence of absorption by greenhouse gases, and the location of spectral absorption bands and window regions – may appreciably alter Earth's "no-feedback" climate sensitivity. These properties could differ considerably for climates of other worlds with much different atmospheric composition – as highlighted in Figure 4. They could also differ markedly for past climates, and they may change to a minor but notable extent as humans continue to alter Earth's atmospheric composition.

The stratospheric masking deviation depends on stratospheric emissivity, which can be altered by anthropogenic greenhouse gases – including effects on stratospheric H_2O from CH_4 oxidation. Using the same reference temperature profile as in Figure 1, a doubling of CO_2, stratospheric H_2O, and O_3, lead to respective increases in Δ_S by 0.05, 0.03, and 0.04 W m^{-2} K^{-1}. The historical combination of small decreases in global stratospheric O_3, together with increases in CO_2 and stratospheric H_2O (e.g., Solomon et al., 2010), has likely made the Planck feedback more positive due to increasing stratospheric opacity. Although these effects are small in the historical period, they will grow in the future and should be accounted for in feedback analysis of climate model simulations that use CO_2 concentrations many times larger than present values. Stratospheric emissivity also differs appreciably between the two radiation codes that we used, raising questions about whether any systematic biases exist in stratospheric opacity in climate models.

An increasingly popular choice, introduced by Held and Shell (2012) (and taken as the default by Zelinka et al., 2020), is to use relative humidity (rather than specific humidity) as a state variable in feedback analysis. Thus, the Planck feedback is calculated with vertically uniform tropospheric warming and constant relative humidity rather than constant specific humidity. This choice then incorporates much of the water vapor feedback into the Planck feedback, which can be an appealing simplification if changes in relative humidity with climate warming or cooling are small (as most models suggest), and also has the benefit of explicitly making the Planck feedback small in spectral regions that are optically thick to water vapor (Jeevanjee et al., 2021). Furthermore, computing the lapse-rate feedback with relative humidity fixed greatly reduces its magnitude, and eliminates the largely cancelling negative covariance across models in the (conventionally-defined) water vapor and lapse-rate feedbacks. Held and Shell (2012) note of this approach in closing that "a drawback is that one loses contact between the reference, 'no feedback,' sensitivity and the simplest textbook estimate based on Stefan-Boltzmann." Our work above suggests that this contact is looser to begin with than has been commonly assumed, and thus breaking contact is perhaps not such a loss. A flip side of this point, however, is that not all of the difference between this "simplest textbook estimate based on Stefan-Boltzmann", $-4\sigma T_e^3$, and the Planck feedback computed with relative humidity fixed, should be blindly attributed to water vapor feedbacks. All the

effects discussed above – stratospheric masking, temperature-dependent opacity, and nonlinear averaging – will still be embedded in the Planck feedback if it is computed with fixed relative humidity.

This paper has explored a subject that most climate scientists would likely consider so simple as to be trivial, and found hidden subtleties in the Planck feedback. These subtleties should be considered explicitly at the level of detail for which feedback analysis is conducted in climate models, and suggest reconsidering the convention of no stratospheric temperature change used in GCM calculations of the Planck feedback. In addition to providing a novel discussion of how the Planck feedback deviates from the negative feedback of a blackbody, this work also serves as a reminder that in the climate system, few things are as simple as they may seem.

APPENDIX A: DECOMPOSING SOURCES OF PATTERN COVARIANCE

To identify the sources of pattern covariance between warming and local Planck feedbacks, we can compute global-average feedbacks that include the structure of both quantities across some subset of the three dimensions (latitude, longitude, time) but not others. We first define the local warming ratio $r(y,\theta,t) = \Delta T(y,\theta,t)/\overline{\Delta T}^{(y,\theta,t)}$ as the ratio of local temperature change $\Delta T(y,\theta,t)$ – dependent on the sine of latitude ($y = \sin\phi$), the longitude θ, and the month of year t – to the global and annual-mean warming $\overline{\Delta T}^{(y,\theta,t)}$. Parentheses following the overline symbol here indicate the dimensions over which the average is taken; e.g., $\overline{\Delta T}^{(\theta)}$ would be the average of the temperature change over longitude, retaining dependence on latitude and time of year. Taking all combinations gives eight different global-average Planck feedbacks which either include or neglect covariance between the warming pattern and the Planck feedback over each of the three dimensions:

$$\begin{aligned}
\overline{\lambda_P}^{(1,1,1)} &= \overline{\lambda_P r}^{(y,\theta,t)} \\
\overline{\lambda_P}^{(1,1,0)} &= \overline{\overline{\lambda_P}^{(t)}\,\overline{r}^{(t)}}^{(y,\theta)} \\
\overline{\lambda_P}^{(1,0,1)} &= \overline{\overline{\lambda_P}^{(\theta)}\,\overline{r}^{(\theta)}}^{(y,t)} \\
\overline{\lambda_P}^{(0,1,1)} &= \overline{\overline{\lambda_P}^{(y)}\,\overline{r}^{(y)}}^{(\theta,t)} \\
\overline{\lambda_P}^{(1,0,0)} &= \overline{\overline{\lambda_P}^{(\theta,t)}\,\overline{r}^{(\theta,t)}}^{(y)} \\
\overline{\lambda_P}^{(0,1,0)} &= \overline{\overline{\lambda_P}^{(y,t)}\,\overline{r}^{(y,t)}}^{(\theta)} \\
\overline{\lambda_P}^{(0,0,1)} &= \overline{\overline{\lambda_P}^{(y,\theta)}\,\overline{r}^{(y,\theta)}}^{(t)} \\
\overline{\lambda_P}^{(0,0,0)} &= \overline{\lambda_P}^{(y,\theta,t)}\,\overline{r}^{(y,\theta,t)}.
\end{aligned} \tag{8}$$

Note here that $\overline{\lambda_P}^{(0,0,0)}$ is the simple average previously defined, and $\overline{\lambda_P}^{(1,1,1)}$ is the full temperature-change-weighted average including covariance of Planck feedback and warming over latitude, longitude, and time of year. Subtracting $\overline{\lambda_P}^{(0,0,0)}$ from each gives a pattern covariance deviation $\Delta_C^{(\cdot,\cdot,\cdot)} = \overline{\lambda_P}^{(\cdot,\cdot,\cdot)} - \overline{\lambda_P}^{(0,0,0)}$, with values for each term in each model shown in Table A1. Comparing the calculated values in Table A1 shows that latitudinal structure dominates the covariance of the warming pattern and local Planck feedback. Temporal structure matters as well, but only when included together with latitudinal structure, and longitudinal structure is almost entirely negligible for the pattern covariance of warming and Planck feedback. To first order, one can think of the pattern covariance deviation term as

arising because warming and Planck feedbacks are larger at the poles than in the tropics, and to second order because both are also larger at the poles in (local) winter than in summer.

APPENDIX B: ON THE NONLINEAR AVERAGING TERMS

The nonlinear averaging deviation term calculated above, Δ_N, does not distinguish between the relative importance of nonlinearity in averaging over heights, as opposed to averaging over wavenumbers. Following logic similar to that in section 2.2, an approximate power-law form of the monochromatic flux can be used to understand why the effect of nonlinear averaging over heights is generally small. The Planck function,

$$B^{\tilde{\nu}}(T) = \frac{2\pi h c^2 \tilde{\nu}^3}{\exp\left(\frac{hc\tilde{\nu}}{kT}\right) - 1}, \tag{9}$$

which describes the radiance of a blackbody at temperature T (where h is Planck's constant, c is the speed of light, and k is Boltzmann's constant), can be approximated near a reference temperature T_0 by:

$$B^{\tilde{\nu}}(T) = B^{\tilde{\nu}}(T_0)\left(\frac{T}{T_0}\right)^{\alpha} \tag{10}$$

$$\alpha \equiv \frac{d\log B^{\tilde{\nu}}(T_0)}{d\log T} = \frac{hc\tilde{\nu}}{kT_0}\frac{\exp\left(\frac{hc\tilde{\nu}}{kT_0}\right)}{\exp\left(\frac{hc\tilde{\nu}}{kT_0}\right) - 1}. \tag{11}$$

(e.g., Jeevanjee and Fueglistaler, 2020). This exponent, α, approaches 1 for photons with much less than average thermal energies ($hc\tilde{\nu}/(kT_0) << 1$), and asymptotes from above to $hc\tilde{\nu}/(kT_0)$ for photons with much greater than average thermal energies ($hc\tilde{\nu}/(kT_0) >> 1$). Denoting the arithmetic mean of emitting temperatures at a single wavenumber as $\overline{T}$, and departures from this emitting temperature as T', it can be shown that:

$$\overline{\frac{dF}{dT}} = \left[\frac{b\pi B^{\tilde{\nu}}(T_0)}{T_0^{\alpha}}\right]\overline{T^{\alpha-1}} \approx \left[\frac{\alpha\pi B^{\tilde{\nu}}(T_0)}{T_0}\right]\overline{T}^{\alpha-1}\left(1 + \frac{(\alpha-1)(\alpha-2)}{2}\frac{\overline{(T')^2}}{\overline{T}^2}\right) \tag{12}$$

$$\left(\frac{dF}{dT}\right)_{T(\overline{F})} = \left[\frac{\alpha\pi B^{\tilde{\nu}}(T_0)}{T_0^{\alpha}}\right]\left(\overline{T^{\alpha}}\right)^{1-\frac{1}{\alpha}} \approx \left[\frac{\alpha\pi B^{\tilde{\nu}}(T_0)}{T_0}\right]\overline{T}^{\alpha-1}\left(1 + \frac{(\alpha-1)^2}{2}\frac{\overline{(T')^2}}{\overline{T}^2}\right), \tag{13}$$

where the latter approximation of each expression has used a Taylor expansion and retained only the variance of emitting temperatures but not higher-order terms, by assuming that $T'/\overline{T}$ is small. Even before each approximation, it should be clear that $\overline{dF}/dT < (dF/dT)_{T(\overline{F})}$, because $\alpha > 1$ and thus $(dF/dT)_{T(\overline{F})}$ is a concave function of the flux derivatives involved in $\overline{dF}/dT$ (e.g., with $\alpha = 4$, $(dF/dT)_{T(\overline{F})}$ is the 3/4 power of an average of the 4/3 power of the flux derivatives used to calculate $\overline{dF}/dT$). Using these approximate forms, the difference between the two terms is given by:

$$\left(\frac{dF}{dT}\right)_{T(\overline{F})} - \overline{\frac{dF}{dT}} \approx \left[\frac{\alpha\pi B^{\tilde{\nu}}(T_0)}{T_0}\right]\overline{T}^{\alpha-1}\left[\frac{\alpha-1}{2}\frac{\overline{(T')^2}}{\overline{T}^2}\right]. \tag{14}$$

In a relative sense, this difference is very close to $\frac{\alpha-1}{2}(\overline{(T')^2}/\overline{T}^2)$, and thus depends on the variance of the emitting temperatures relative to the square of the mean emitting temperature. Note that the result from section 2.2 is a specific example of this difference that can be recovered by using $\alpha = 4$ and $(\overline{(F')^2}/\overline{F}^2) = 16(\overline{(T')^2}/\overline{T}^2)$. We can roughly estimate a typical relative deviation from the mean emitting temperature based on how rapidly the logarithm of atmospheric temperature changes with the logarithm of monochromatic optical thickness (using the reasoning that emission mostly occurs within one factor of e near $\tau = 1$). Using the chain rule to express $d \log T / d \log \tau$ gives:

$$\frac{|T'|}{\overline{T}} \approx \frac{d \log T}{d \log \tau} = \frac{d \log T}{d \log p} \times \frac{d \log p}{d \log \tau} \approx \frac{R\Gamma}{g} \frac{1}{\beta}, \tag{15}$$

where Γ is the lapse rate, and use of $\beta = d \log \tau / d \log p$ follows Jeevanjee and Fueglistaler (2020). Applying this expression for a few examples shows why the deviation of $(dF/dT)_{T(\overline{F})}$ from $\overline{dF/dT}$ is typically small. For a well-mixed gas such as CO_2, $\beta = 2$ due to pressure broadening; if emission is mostly tropospheric, we can take $\Gamma \approx 6.5$ K km^{-1} to obtain $|T'|/\overline{T} \approx 0.1$, and thus a nonlinear averaging deviation on the order of 1-2% from averaging across heights. For wavenumbers where CO_2 emits from the stratosphere, however, the magnitude of the lapse rate is greatly reduced and thus the deviation is smaller by about a factor of 10. For H_2O emitting from the troposphere, the larger value of $\beta \approx 5$ due to decay in H_2O concentration with height leads to expected deviations on the order of 0.1% (and made further smaller by the low values of α in the H_2O rotational band). Note that the key parameter for the relative magnitude of this deviation term, $\frac{\alpha-1}{2}(\overline{(T')^2}/\overline{T}^2)$, can also be written as $\frac{\alpha-1}{2\alpha^2}\gamma^2$, where $\gamma = d \log B / d \log \tau$ is the key parameter governing the validity of the cooling-to-space approximation (valid for $\gamma << 1$), derived by Jeevanjee and Fueglistaler (2020). The connection between neglect of the within-wavenumber nonlinear averaging term and the validity of the cooling-to-space approximation arises because both conditions hold best when the levels that emit to space have only a small variation in temperature (leading to small variations in the Planck source function).

One plausible situation where nonlinear averaging across heights could be more important would be the case of a high cloud of optical thickness ~1 near the tropopause, in a spectral region with very low clear-sky optical thickness and a large value of α: in such conditions averaging across heights might introduce a deviation as large as 10% of the actual monochromatic flux change. This caveat noted, we conclude that nonlinear averaging across heights is generally a small deviation term, so Δ_N is mostly determined by nonlinear averaging across wavenumbers.

ACKNOWLEDGEMENTS

Work on this paper was supported by NSF Atmospheric Chemistry grant AGS-1906719: "Advancing the Understanding of the Impacts of Wave-Induced Temperature Fluctuations on Atmospheric Chemistry". We thank Susan Solomon, Peter Molnar, Nick Lutsko, Daniel Koll, and Aaron Donohoe for helpful comments on drafts of this manuscript, Nadir Jeevanjee and two anonymous reviewers for useful feedback (particular thanks is due to the reviewer who penned two lengthy and rigorous reviews that heavily reshaped the paper and led to the coinage of "Stefan-Boltzmann feedback"), and AER for freely providing LBLRTM (available at `http://rtweb.aer.com/lblrtm.html`). We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table A1 of this paper) for producing and making available their model output. Model output and scripts used to make figures in this paper are available at `https://www.dropbox.com/sh/9vrn1v4brxj2m7m/AACp1jELMc1OfLWo3ZoBDwsma?dl=0`.

REFERENCES

Back, L., Russ, K., Liu, Z., Inoue, K., Zhang, J. and Otto-Bliesner, B. (2013) Global hydrological cycle response to rapid and slow global warming. *Journal of Climate*, **26**, 8781–8786.

Bader, D. C., Leung, R., Taylor, M. and McCoy, R. B. (2019) E3SM-Project E3SM1.0 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Bethke, I., Wang, Y., Counillon, F., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H. R., Chiu, P.-G., Bentsen, M., Guo, C., Tjiputra, J., Kirkevag, A., Olivie, D. J. L., Seland, O., Fan, Y., Lawrence, P., Eldevik, T. and Keenlyside, N. (2019) NCC NorCPM1 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Bony, S., Colman, R., Kattsov, V. M., Allan, R. P., Bretherton, C. S., Dufresne, J.-L., Hall, A., Hallegatte, S., Holland, M. M., Ingram, W., Randall, D. A., Soden, B. J., Tselioudis, G. and Webb, M. J. (2006) How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, **19**, 3445–3482.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Balkanski, Y., Checa-Garcia, R., Hauglustaine, D., Bekki, S. and Marchand, M. (2021) IPSL IPSL-CM6A-LR-INCA model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Byun, Y.-H., Lim, Y.-J., Sung, H. M., Kim, J., Sun, M. and Kim, B.-H. (2019) NIMS-KMA KACE1.0-G model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Cao, J. and Wang, B. (2019) NUIST NESMv3 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S. and Brown, P. D. (2005) Atmospheric radiative transfer modeling: A summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer*, **91**, 233–244.

Colman, R. (2003) A comparison of climate feedbacks in general circulation models. *Climate Dynamics*, **20**, 865–873.

Cormier, J. G., Hodges, J. T. and Drummond, J. R. (2005) Infrared water vapor continuum absorption at atmospheric temperatures. *The Journal of Chemical Physics*, **122**.

Danabasoglu, G. (2019) NCAR CESM2-WACCM model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, Ian Law, R., Mackallah, C., Marsland, S., O'Farrell, S., Rashid, H., Srbinovsky, J., Sullivan, A., Trenham, C., Vohralik, P., Watterson, I., Williams, G., Woodhouse, M., Bodman, R., Dias, F. B., Domingues, C., Hannah, N., Heerdegen, A., Savita, A., Wales, S., Allen, C., Druken, K., Evans, B., Richards, C., Ridzwan, S. M., Roberts, D., Smillie, J., Snow, K., Ward, M. and Yang, R. (2019) CSIRO-ARCCSS ACCESS-CM2 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Feldl, N. and Roe, G. H. (2013) The nonlinear and nonlocal nature of climate feedbacks. *Journal of Climate*, **26**, 8289–8304.

Fu, Q. and Liou, K. N. (1992) On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *Journal of the Atmospheric Sciences*, **49**, 2139–2156.

Guo, H., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Bushuk, M., Dunne, K. A., Dussin, R., Gauthier, P. P., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Lin, P., Malyshev, S., Naik, V., Paulot, F., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Shao, A., Silvers, L., Wyman, B., Yan, X., Zeng, Y., Adcroft, A., Dunne, J. P., Held, I. M., Krasting, J. P., Horowitz, L. W., Milly, P., Shevliakova, E., Winton, M., Zhao, M. and Zhang, R. (2018) NOAA-GFDL GFDL-CM4 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Hajima, T., Abe, M., Arakawa, O., Suzuki, T., Komuro, Y., Ogura, T., Ogochi, K., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Ohgaito, R., Ito, A., Yamazaki, D., Ito, A., Takata, K., Watanabe, S., Kawamiya, M. and Tachiiri, K. (2019) MIROC MIROC-ES2L model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Hansen, J., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., Ruedy, R. and Lerner, J. (1984) Climate sensitivity: Analysis of feedback mechanisms. *Washington DC American Geophysical Union Geophysical Monograph Series*, **29**, 130–163.

Hansen, J., Sato, M. and Ruedy, R. (1997) Radiative forcing and climate response. *Journal of Geophysical Research*, **102**, 6831–6864.

Held, I. M. and Shell, K. M. (2012) Using relative humidity as a state variable in climate feedback analysis. *Journal of Climate*, **25**, 2578–2582.

Huang, Y. and Ramaswamy, V. (2007) Effect of the temperature dependence of gas absorption in climate feedback. *Journal of Geophysical Research: Atmospheres*, **112**.

Ingram, W. (2010) A very simple model for the water vapour feedback on climate change. *Quarterly Journal of the Royal Meteorological Society*, **136**, 30–40.

Jeevanjee, N. and Fueglistaler, S. (2020) On the cooling to space approximation. *Journal of the Atmospheric Sciences*, **77**, 465–478.

Jeevanjee, N., Koll, D. D. B. and Lutsko, N. (2021) "Simpson's Law" and the spectral cancellation of climate feedbacks. *Geophysical Research Letters*, **48**.

Kim, Y. H., Noh, Y., Kim, D., Lee, M.-I., Lee, H. J., Kim, S. Y. and Kim, D. (2019) KIOST KIOST-ESM model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G. and Kato, S. (2018) Clouds and the earth's radiant energy system (ceres) energy balanced and filled (ebaf) top-of-atmosphere (toa) edition-4.0 data product. *Journal of Climate*, **31**, 895–918.

NASA/GISS (2018) NASA-GISS GISS-E2.1G model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

— (2019) NASA-GISS GISS-E2.1H model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Oman, L., Waugh, D. W., Pawson, S., Stolarski, R. S. and Nielsen, J. E. (2008) Understanding the changes of stratospheric water vapor in coupled chemistry-climate model simulations. *Journal of the Atmospheric Sciences*, **65**, 3278–3291.

Oreopoulos, L., Mlawer, E., Delamere, J., Shippert, T., Cole, J., Fomin, B., Iacono, M., Jin, Z., Li, J., Manners, J., Raisanan, P., Rose, F., Zhang, Y., Wilson, M. J. and Rossow, W. B. (2012) The continual intercomparison of radiation codes: Results from phase I. *Journal of Geophysical Research: Atmospheres*, **117**.

Shine, K. P., Ptashnik, I. V. and Radel, G. (2012) The water vapor continuum: brief history and recent developments. *Surveys in Geophysics*, **33**, 535–555.

Simpson, G. C. (1928) Some studies in terrestrial radiation. *Memoirs of the Royal Meteorological Society*, **2**, 69–95.

Soden, B. J. and Held, I. M. (2006) An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate*, **19**, 3354–3360.

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

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J. and Plattner, G.-K. (2010) Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science*, **327**, 1219–1223.

Stouffer, R. (2019) UA MCM-UA-1-0 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Solheim, L., von Salzen, K., Yang, D., Winter, B. and Sigmond, M. (2019) CCCma CanESM5 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Tatebe, H. and Watanabe, M. (2018) MIROC MIROC6 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Wang, Y. and Huang, Y. (2020) Stratospheric radiative feedback limited by the tropospheric influence in global warming. *Climate Dynamics*, **55**, 2343–2350.

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S. and Ohno, T. (2018) Radiative-convective equilibrium model intercomparison project. *Geoscientific Model Development*, **11**, 793–813.

Wu, T., Chu, M., Dong, M., Fang, Y., Jie, W., Li, J., Li, W., Liu, Q., Shi, X., Xin, X., Yan, J., Zhang, F., Zhang, J., Zhang, L. and Zhang, Y. (2018) BCC BCC-CSM2MR model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Yu, Y. (2019) CAS FGOALS-f3-L model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yoshimura, H., Shindo, E., Mizuta, R., Ishii, M., Obata, A. and Adachi, Y. (2019) MRI MRI-ESM2.0 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

Zelinka, M. D., Meyers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A. and Taylor, K. E. (2020) Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, **47**.

Zhang, J., Wu, T., Shi, X., Zhang, F., Li, J., Chu, M., Liu, Q., Yan, J., Ma, Q. and Wei, M. (2019) BCC BCC-ESM1 model output prepared for CMIP6; historical, abrupt-4xCO2. *Earth System Grid Federation*.

TABLE 1 CMIP6 models used and decomposition of meridional covariance deviation for each.

Climate Model	Citation	$\Delta_C^{(1,1,1)}$	$\Delta_C^{(1,1,0)}$	$\Delta_C^{(1,0,1)}$	$\Delta_C^{(0,1,1)}$	$\Delta_C^{(1,0,0)}$	$\Delta_C^{(0,1,0)}$	$\Delta_C^{(0,0,1)}$
		$W\,m^{-2}\,K^{-1}$						
ACCESS-CM2	Dix et al. (2019)	0.071	0.063	0.080	-0.002	0.072	-0.002	0.001
BCC-CSM2-MR	Wu et al. (2018)	0.088	0.073	0.098	0.000	0.083	-0.001	0.002
BCC-ESM1	Zhang et al. (2019)	0.093	0.080	0.102	0.001	0.088	0.000	0.002
CanESM5	Swart et al. (2019)	0.081	0.069	0.086	0.002	0.074	0.001	0.001
CESM2	Danabasoglu (2019)	0.074	0.062	0.086	-0.001	0.075	-0.001	0.001
E3SM1.0	Bader et al. (2019)	0.080	0.074	0.088	-0.001	0.082	-0.001	0.001
FGOALS-f3-L	Yu (2019)	0.103	0.085	0.113	0.000	0.094	-0.002	0.002
GFDL-CM4	Guo et al. (2018)	0.072	0.065	0.081	-0.002	0.074	-0.002	0.000
GISS-E2.1G	NASA/GISS (2018)	0.052	0.042	0.058	-0.001	0.048	0.000	0.001
GISS-E2.1H	NASA/GISS (2019)	0.080	0.067	0.093	-0.004	0.080	-0.005	0.001
IPSL-CM6A-LR-INCA	Boucher et al. (2021)	0.072	0.066	0.078	0.000	0.072	0.001	0.001
KACE1.0-G	Byun et al. (2019)	0.072	0.062	0.080	-0.001	0.071	-0.001	0.001
KIOST-ESM	Kim et al. (2019)	0.042	0.029	0.049	-0.001	0.037	-0.002	0.001
MCM-UA-1-0	Stouffer (2019)	0.084	0.068	0.093	-0.001	0.078	0.000	0.000
MIROC6	Tatebe and Watanabe (2018)	0.049	0.039	0.059	0.001	0.049	0.000	0.002
MIROC-ES2L	Hajima et al. (2019)	0.054	0.043	0.060	0.000	0.049	-0.001	0.001
MRI-ESM2.0	Yukimoto et al. (2019)	0.099	0.084	0.107	-0.002	0.093	-0.002	0.000
NESMv3	Cao and Wang (2019)	0.095	0.085	0.099	0.001	0.090	0.000	0.001
NorCPM1	Bethke et al. (2019)	0.124	0.111	0.128	0.001	0.115	0.001	0.001
Multi-model mean		0.078	0.067	0.086	-0.001	0.075	-0.001	0.001
Multi-model std. dev.		0.020	0.019	0.020	0.001	0.018	0.001	0.001

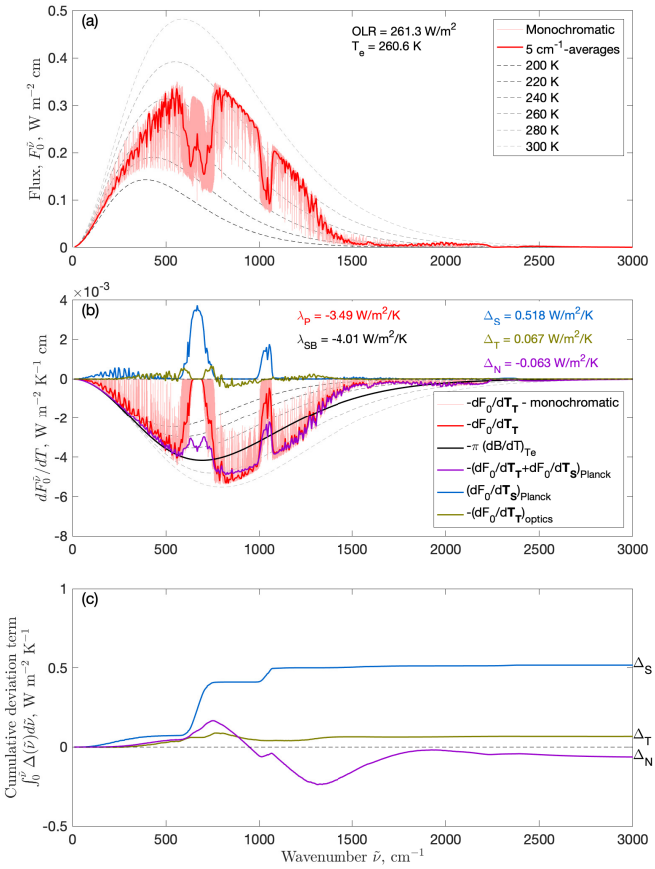


FIGURE 1 a) Outgoing infrared flux spectrum with the line-by-line radiative transfer model LBLRTM, for an atmosphere with a surface temperature of 290 K. The pink curve indicates monochromatic irradiances, and the red curve irradiances averaged over 5 cm^{-1} wavenumber bands. Thin dashed lines from black to light gray show reference blackbody spectra. b) Changes in radiative flux spectra for: 1K of troposphere and surface cooling (monochromatic change in pink, average over 5 cm^{-1} bands in red), a 1K decrease in the temperature of a blackbody at temperature T_e (black), a 1K decrease of the temperature for the Planck source function at all levels (purple), a 1K increase in temperature for the Planck source function in the stratosphere (blue), and a 1K decrease in gas-optics calculations in the troposphere (gold). Thin dashed lines from black to light gray show negative values of the derivative of the Planck function for the same reference temperatures as in a). c) Cumulative integrals from 0 to wavenumber $\tilde{\nu}$ of each deviation term. The value at the right-hand side of the plot indicates the full spectral integral of the deviation term Δ, and wavenumbers of greatest slope indicate areas most important to the term.

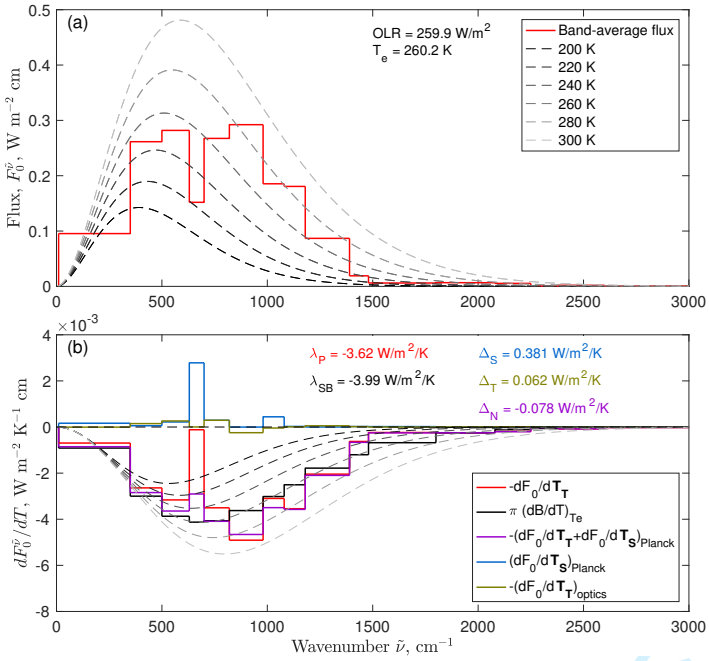


FIGURE 2 a) Outgoing infrared flux spectrum with the correlated-k radiative transfer model RRTMG, for the same atmospheric temperature and trace-gas profiles as used in Figure 1. Thin dashed lines from black to light gray show reference blackbody spectra. b) Changes in OLR flux spectrum for: 1K of troposphere and surface cooling (average over bands in red), a 1K decrease in the temperature of a blackbody at temperature T_e (black), a 1K decrease of the temperature for the Planck source function at all levels (purple), a 1K increase in temperature for the Planck source function in the stratosphere (blue), and a 1K decrease in gas-optics calculations in the troposphere (gold). Thin dashed lines from black to light gray show negative values of the derivative of the Planck function for the same reference temperatures as in a). Total values of each deviation term are included in b).

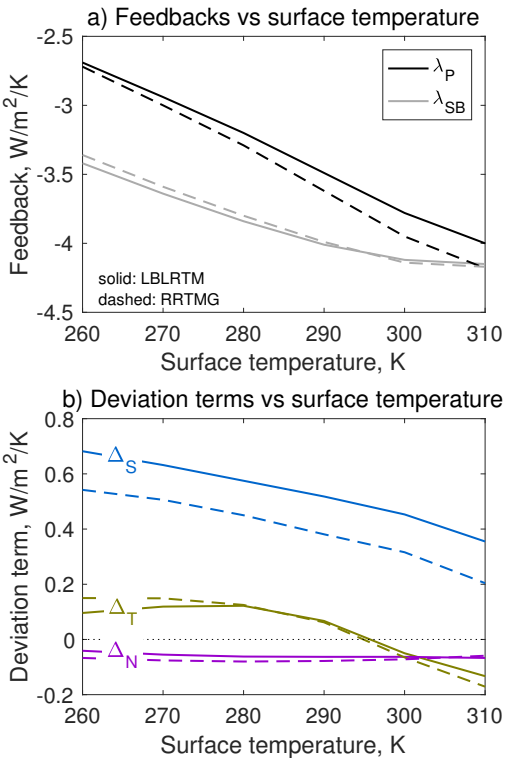


FIGURE 3 a) Planck λ_P and Stefan-Boltzmann λ_{SB} feedbacks over a range of surface temperatures, for calculations with LBLRTM (solid) and RRTMG (dashed). b) Spectrally integrated deviation terms for stratospheric masking (Δ_S), temperature-dependent opacity (Δ_T), and nonlinear averaging (Δ_N) over a range of surface temperatures, for calculations with LBLRTM (solid) and RRTMG (dashed).

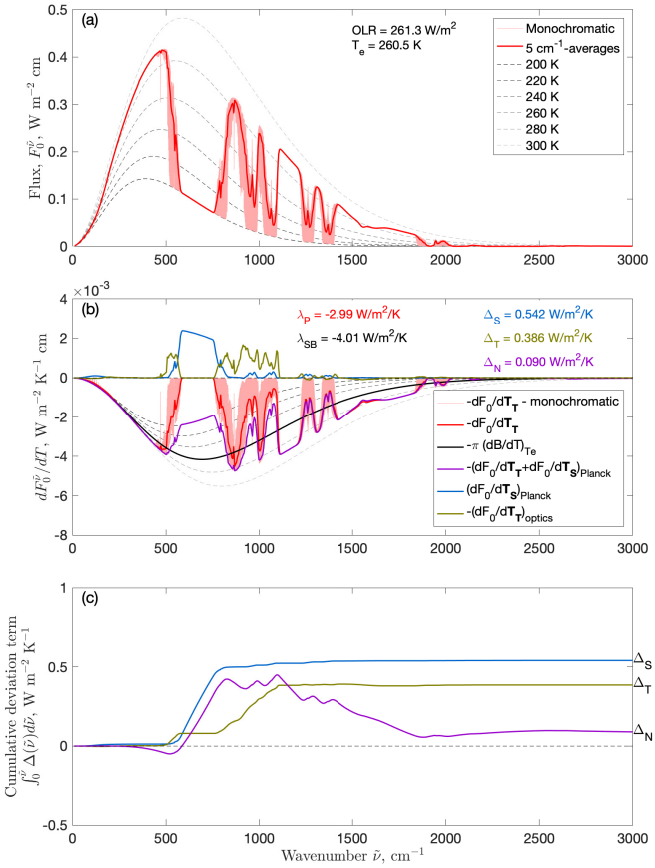


FIGURE 4 As in Figure 1 but for an atmosphere with no H_2O and instead 19.5% CO_2; the temperature profile follows a dry adiabat in the troposphere up to an isothermal stratosphere at 200 K. a) Outgoing infrared flux spectrum; b) Changes in OLR flux spectrum for different cases and spectra of different deviation terms. c) Cumulative integrals from 0 to wavenumber $\tilde{\nu}$ of each deviation term.