



A formal analysis of the feedback concept in climate models. Part I: Exclusive and inclusive feedback analyses

Alain Lahellec, Jean-Louis Dufresne

► To cite this version:

Alain Lahellec, Jean-Louis Dufresne. A formal analysis of the feedback concept in climate models. Part I: Exclusive and inclusive feedback analyses. *Journal of the Atmospheric Sciences*, 2013, 70 (12), pp.3940-3958. 10.1175/JAS-D-12-0218.1 . hal-01089499v2

HAL Id: hal-01089499

<https://hal.science/hal-01089499v2>

Submitted on 2 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Formal Analysis of the Feedback Concept in Climate Models. Part I: Exclusive and Inclusive Feedback Analyses*

ALAIN LAHELLEC AND JEAN-LOUIS DUFRESNE

Laboratoire de Météorologie Dynamique, Paris, France

(Manuscript received 30 July 2012, in final form 24 July 2013)

ABSTRACT

Climate sensitivity and feedback are key concepts if the complex behavior of climate response to perturbation is to be interpreted in a simple way. They have also become an essential tool for comparing global circulation models and assessing the reason for the spread in their results. The authors introduce a formal basic model to analyze the practical methods used to infer climate feedbacks and sensitivity from GCMs. The tangent linear model is used first to critically review the standard methods of feedback analyses that have been used in the GCM community for 40 years now. This leads the authors to distinguish between exclusive feedback analyses as in the partial radiative perturbation approach and inclusive analyses as in the “feedback suppression” methods. This review explains the hypotheses needed to apply these methods with confidence. Attention is paid to the more recent regression technique applied to the abrupt $2\times\text{CO}_2$ experiment. A numerical evaluation of it is given, related to the Lyapunov analysis of the dynamical feature of the regression. It is applied to the Planck response, determined in its most strict definition within the GCM. In this approach, the Planck feedback becomes a dynamical feedback among others and, as such, also has a fast response differing from its steady-state profile.

1. Introduction

The studies on climate sensitivity initiated by Manabe and Wetherald (1975), which apply general circulation models, spectacularly complete Charney’s program. By already pointing out that water vapor was the main driver for the amplification of climate response to an increase in CO_2 concentration, the earliest studies (Manabe and Wetherald 1975; Coakley 1977; Hansen et al. 1984; Wetherald and Manabe 1988) gave rise to numerical simulations to answer questions that had been raised by Fourier and Arrhénius (1896) in the nineteenth century: What determines the atmospheric temperature of Earth? What are the main mechanisms accounting for climate response to perturbations?

The start of the industrial era and the subsequent release of greenhouse gases into the atmosphere presented

a true application case for issues that had remained theoretical until then. Global warming was raised to a “social issue” by the Intergovernmental Panel on Climate Change (IPCC), which initiated numerous scientific investigations with the support of the United Nations Environment Programme (UNEP) and the World Meteorological Organization (WMO).

As a general result, a common understanding of the global warming mechanism can be summarized as follows: an increase in CO_2 concentration directly brings about an increase in atmospheric temperature due to the greenhouse effect produced by this gas. Climate response to an increase in temperature is the result of a chain of numerous complex meteorological mechanisms. One goal of climate analysis—which includes the feedback analysis—is to synthesize these mechanisms and gain an understanding of their complexity.

It is commonly agreed that the global analysis of annually averaged climatic variables delivers the essential characteristics of climate change, measured in terms of a temperature-change ΔT_s in the global surface air temperature. If the climate reaches a new equilibrium, the top-of-the-atmosphere (TOA) radiation budget is restored to zero, implying that the principal climatic variables—that is, temperature and water vapor (and also clouds, surface albedo, etc.)—have changed. Let us

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JAS-D-12-0218.s1>.

Corresponding author address: Alain Lahellec, LMD/IPSL/ Université Pierre et Marie Curie, boîte 99, 4 Place Jussieu, 75252 Paris CEDEX 05, France.
E-mail: alain@lmd.jussieu.fr

suppose that the doubling of $[\text{CO}_2]$ leads to an increase of T_s and let us assume, for simplicity, no change in albedo. An increase of longwave (LW) radiation follows the increase in temperature and, to enable a return to equilibrium, the surplus in LW emission is necessarily retained in the atmosphere somehow. Slingo et al. (2000) use this feature to evaluate the greenhouse effect as the difference between the surface emission and the TOA budget. Climate feedback analysis investigates which mechanisms are responsible for the consecutive change in the greenhouse effect, eventually allowing a return to equilibrium.

Two modes of analysis were initially proposed to quantify climatic feedbacks: one in terms of the sensitivity of TOA budgets to surface temperature change, the other using the concept of feedback gain, which focuses on the system as damping or amplifying a perturbation of the surface temperature. Both approaches share the hypothesis that the climate responds linearly to a small perturbation. In Bode's linear electric circuits theory (Bode 1945), a feedback is quantified by a feedback gain g . A positive gain characterizes an amplifying system (positive feedback), while a negative gain shows a damping system (negative feedback). Also, a gain is associated with a feedback loop, and the opening of the loop results in the canceling of g : this prevents the reaction of the system from affecting the primary perturbation of T_s .

The other approach, initiated by Coakley (1977) and Wetherald and Manabe (1988), proposed an analysis framework now known as the partial radiative perturbation method (PRP). PRP has become a widely accepted method for numerical climate analysis, as seen from the numerous related publications in the GCM community: Soden et al. (2008), Bony et al. (2006), Colman (2003), Colman et al. (1997), Zhang et al. (1994), LeTreut et al. (1994), and Cess et al. (1989), among others. PRP analyzes the relation between N , the net TOA radiative imbalance (outgoing flux), and ΔT_s —the global variation of the planetary surface air temperature at the perturbed equilibrium. Wetherald and Manabe (1988) write this relation as $N = \Delta_x R - \lambda_T \Delta T_s$, where $\Delta_x R$ is the immediate forcing induced by the doubling of $[\text{CO}_2]$ and λ_T is defined as the total climate “sensitivity factor.”

This approximation allows simple relations to come into play and were given in Schlesinger (1986). In particular, both approaches use the same definition of the “Planck response” ($\Delta\theta_0$ for surface temperature), which is the pure radiative response to the doubling of $[\text{CO}_2]$, or the “no feedback” response. The two approaches thus consider that the climate system responds to the Planck perturbation. A doubling- $[\text{CO}_2]$ experiment results

in an increase ΔT_s of the surface temperature at equilibrium: the forcing, $\Delta_x R$, is positive (the CO_2 greenhouse effect); the Planck sensitivity, λ_0 , is negative; and $\lambda = \lambda_T - \lambda_0$ is found in models to be positive, as is the feedback gain g . However, λ_T is negative, ensuring stability of the perturbed climate: the positive climate feedback only means that the output radiation change induced by the CO_2 greenhouse effect is reduced by the other climate mechanisms, leading to an increase of the surface temperature warming compared to the Planck response. Typical values found with GCMs are $\Delta T_s \simeq 3 \text{ K}$, $\Delta\theta_0 \simeq 1.2 \text{ K}$.

Linearity at the global scale is generally verified in the studies cited; its crucial advantage is that it allows the global climate response to be decomposed into additive elementary feedbacks. More recently, Dufresne and Bony (2008), in parallel with Lu and Cai (2009), proposed further decomposing the global warming ΔT_s into elementary contributions: $\Delta T_s = \Delta\theta_0 + \sum_i \Delta\theta_i$, setting $\Delta\theta_i = g_i \Delta T_s$, where $g_i = -\lambda_i/\lambda_0$ is an elementary feedback gain, thus retrieving the definition formerly proposed by Hansen et al. (1984). The Climate Feedback Response Analysis Model (CFRAM) offline approach proposed by Lu and Cai (2009) gives a physical foundation of this decomposition that we analyze in A. Lahellec and J.-L. Dufresne (2013, unpublished manuscript, hereafter Part II).

A major difference between the PRP analysis of perturbed quantities and the feedback approach is that the latter supposes that the model is built with a specific structure: it requires the identification of a feedback loop. The suppression of a feedback mechanism corresponds to the opening of the loop. This is what the “suppression method” (e.g., Hall 2004; Hall and Manabe 1999; Hansen et al. 1997) intended to do. However, as GCMs do not explicitly exhibit such structures, problems arise concerning the consistency of freezing perturbations and still conserving physical constraints (e.g., Schneider et al. 1999).

It has to be emphasized that the two approaches are of radically different natures, contrary to what is suggested by the existence of simple relations between the two, because the sensitivities λ_i are based on fields of variables, while the feedback analysis focuses on the mechanisms responsible for perturbations.

The next section of the paper recalls the results of the two approaches to feedback analysis and what fundamental characteristics can be rigorously deduced from it. Because we are dealing with perturbations, the tangent linear system (TLS) is a privileged tool for formal analyses. It is introduced in section 3 with the classical feedback definition based on a feedback loop. An original formalism is used to represent the dynamic climate

system in abstract form. Once its basic elements are understood, it will be revealed as the most adequate tool to solve this integro-differential system because it only requires simple matrix algebra. It provides some clarifications about various practices in use in the numerical climate community. In particular, section 4 introduces the inclusive approach and examines the reality of decomposing the global warming into meaningful components. Application to the PRP approach is then established in section 5. Up to this point, the dynamics of the climatic system is only put forward as an extension of the steady-state analyses, whereas it is fully exploited in the following sections. In section 5b, the regression method proposed in Gregory et al. (2004) is critically reviewed, because it ushered in a new era in Coupled Model Intercomparison Project (CMIP) analyses, and deserves a strong mathematical foundation to avoid misapplications, some of which have already occurred. The last section proposes a new definition of the Planck response, introducing a new perturbation algorithm in GCMs. It is used to analyze the problem raised by the dynamical features underlying the regression technique and to illustrate our abstract formal developments.

The article ends with a discussion of the applicability of our formal results.

2. Climate sensitivity and the separation issue

In this section, we summarize how the two approaches to climate sensitivity pragmatically separate the global sensitivity into components.

a. Partial radiative perturbation approach

The PRP approach to climate sensitivity considers the numerical results from GCMs as seen from space. If ΔR_x is a net (incoming) TOA forcing, the climate system reacts with a change in the outgoing radiation ΔR . The net budget is $N = \Delta R_x - \Delta R$, which is equal to zero at the new perturbed equilibrium. For small perturbations, if ΔR is found linearly dependent on change in surface temperature ΔT_s , then this global dependence can be decomposed, at equilibrium, into a sum of contributions attributable to separate origins as follows.

Radiative sensitivity is analyzed in terms of the effect of perturbation profiles $\Delta \mathbf{x}_i$ of all climate variables $\mathbf{x}_i$ directly impacting the TOA radiation budget N : temperatures T , specific humidity Q , etc. The R sensitivity is written as the row matrix $\partial R / \partial x_i(k)$: it is a measure of how strongly each entry (k) of the perturbation vector $\mathbf{x}_i$ impacts R —it can be computed offline using GCM radiation codes only.

Each profile $\Delta \mathbf{x}_i$ is taken from a perturbation experiment as an anomaly with respect to the reference run. If we add the hypothesis that T_s , the surface air temperature, is the source of the perturbation, this becomes $\Delta \mathbf{x}_i(k) = [dx_i(k)/dT_s] \Delta T_s$. In this case, a sensitivity factor is defined as the scalar product—using the Einstein summation rule, here on layers k ,

$$\lambda_i = -\frac{\Delta R_i}{\Delta T_s} = -\frac{\partial R}{\partial x_i(k)} \frac{dx_i(k)}{dT_s}. \quad (1)$$

If the effect of $\Delta \mathbf{x}_i$ is to reduce the outgoing radiation R ($\Delta R_i < 0$), the feedback factor is considered positive feedback.

Because the contributions to R are independent, the sum of all the λ_i gives the total change in the TOA budget:

$$\Delta R = -\left(\lambda_0 + \sum_i \lambda_i\right) \Delta T_s. \quad (2)$$

The first term inside the parentheses is the so-called Planck feedback factor λ_0 —the others represent climatic factors. The Planck response, as defined in Soden et al. (2004) for example, is a pure physical effect driven directly by the temperature (blackbody radiation): any source of increase in the atmospheric temperature will result in an increase in the outgoing longwave radiation (OLR): λ_0 is negative. This is taken as a basic response that initiates a change in meteorological mechanisms, which in turn are considered as the true climatic response.

In Dufresne and Bony (2008) for example, the global feedback factor is decomposed as

$$\lambda = \sum_i \lambda_i = \lambda_W + \lambda_\Gamma + \lambda_C + \lambda_\alpha, \quad (3)$$

the sum over partial factors associated with the water-vapor feedback and the lapse rate and clouds and surface-albedo feedbacks—a now classical decomposition. In this decomposition, it is implied that the original perturbation directly impacted T_s —leaving the $\Delta \mathbf{x}_i$ as only perturbed by ΔT_s .

b. Link with feedback gains

As established in Schlesinger (1986), the link with feedback gains can be seen when both sides of Eq. (2) are divided by $-\lambda_0$:

$$\left(1 - \sum_i \frac{\overbrace{-\lambda_i}^{g_i}}{\lambda_0}\right) \Delta T_s = -\frac{\overbrace{\Delta R}^{\Delta \theta_0}}{\lambda_0}. \quad (4)$$

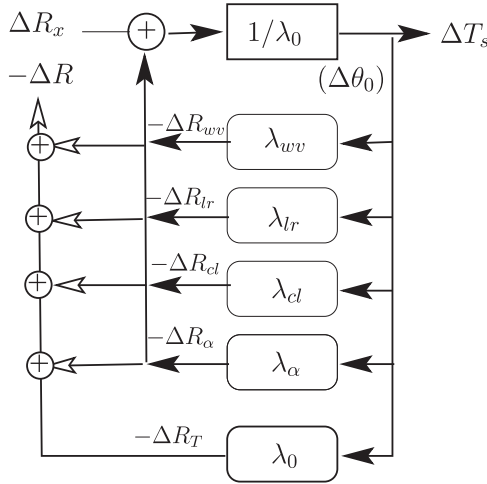


FIG. 1. PRP feedback structure. The basic forcing $\Delta\theta_0$ excites four feedbacks in parallel explaining the global temperature increase (black arrows). Seen from space, each feedback contributes with the Planck response to the radiation budget (white arrows).

This is the classic form of the feedback gain equation and it defines the total response ΔT_s of the system subjected to the forcing perturbation $\Delta\theta_0$. The rhs term $\Delta\theta_0 = -\Delta R/\lambda_0$ is the initial perturbation (i.e., the response of the system when R is insensitive to the feedback mechanisms). It gives the corresponding Planck response $\Delta\theta_0$.

Because of the additivity of the λ_i , the gains add up, giving

$$\left(1 - \sum_i g_i\right) \Delta T_s = \Delta\theta_0, \quad (5)$$

thus satisfying the rule for the addition of feedback gains as in Pedlosky (1982) for instance.

So, apart from the very particular role played by the Planck response, whose definition, as we will see, is not straightforward, as soon as global perturbation profiles are known, any decomposition that sums up elementary terms leads to the definition of feedback gains. Because the λ_i all add up, it follows that the gains $g_i = -\lambda_i/\lambda_0$ are also cumulative. It furthermore follows that the temperature decomposition $\Delta T_s = \Delta\theta_0 + \sum_i \Delta\theta_i$ is also possible; it is sufficient to adopt $\Delta\theta_i = g_i \Delta T_s$. However, the interpretation in terms of feedback loop is not straightforward.

c. Feedback analysis issue

In Fig. 1, the diagram represents the PRP interpretation of the effect of a forcing ΔR_x on the system. When no feedback is active, the response is $\Delta\theta_0$. This forcing of

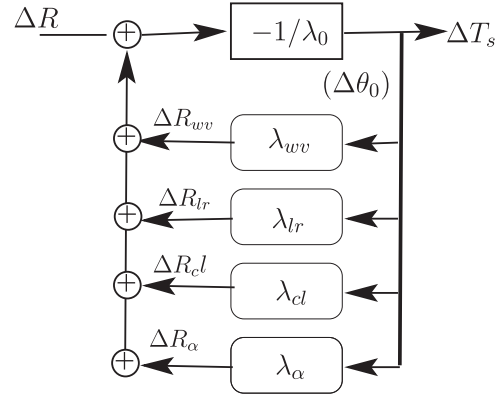


FIG. 2. PRP feedback structure at equilibrium. Each feedback agent adds an internally generated specific forcing to the primary forcing.

the surface temperature excites different mechanisms associated with the feedback factors λ_i , which in turn influence the TOA flux. To reach equilibrium, the temperature factor effect $\Delta R_T = -\lambda_0 \Delta T_s$ on the TOA budget is added to the feedback radiative contributions ΔR_i , so that $\Delta R = -\lambda_0 \Delta T_s - \sum \lambda_i \Delta T_s = \Delta R_T + \sum \Delta R_i$.

At equilibrium, $\Delta R_x - \Delta R = 0$, and the feedback structure simplifies to the one in Fig. 2.

Current formulas connected with the PRP are

$$\Delta R = -\lambda_0 \Delta\theta_0 = -\lambda_T \Delta T_s,$$

$$\Delta T_s = \Delta\theta_0 + \sum_i \Delta\theta_i, \quad \text{and}$$

$$\Delta\theta_i = -\frac{\lambda_i}{\lambda_0} \Delta T_s = -\frac{\lambda_i}{\lambda_T} \Delta\theta_0, \quad (6)$$

where $\lambda_T = \lambda_0 + \lambda$. The decomposition of the climatic response into feedback mechanisms is seen from the impact of ΔT_s on ΔR . If λ_i is to be associated with a climatic mechanism, so is $\Delta\mathbf{x}_i$: the elementary perturbations should be the result of some well-defined climatic process. In that case, the elementary feedback loops act in parallel, each characterized by its feedback gain. Thus we should be able to identify the part that is responsible for this gain in a model. This is one of the goals this article is intended to attain.

However, it needs to be emphasized, as in Ingram et al. (1989), that the gains thus defined are very sensitive to the choice of the basic response: We have $dg/g = [(1-g)/g]d\lambda_0/\lambda_0$ and it follows that, for moderate gain ($g \simeq 50\%$), an imprecision in the determination of λ_0 propagates to the gain value.

Note, also, that the procedure is based on a direct perturbation of T_s , while, in reality, the climate is

perturbed in the bulk of the whole atmosphere, whether the forcing is due to the shortwave (SW) input or stems from a doubling of $[\text{CO}_2]$; this is a reason why we need a more detailed formalism than the classical approach.

3. Sensitivity and feedback in climate models

The “feedback” is historically introduced as measured by the feedback gain. We first formalize this approach before dealing with the more recent climate sensitivity approach. Originally introduced in linear electric systems analysis by Bode (1945) to quantify the importance of feedback mechanisms, the feedback gain g is defined by an equation of the form

$$(1 - g)\Delta T_s = \varpi. \quad (7)$$

The simplicity of this analysis stems from the fact that a scalar variable is extracted from the full system as a “test variable”—the only variable that is originally perturbed (by ϖ) and considered for analysis. The question is how this simplicity can be preserved in the more complex context of climate feedback analysis.

Consider the atmospheric state-space model driven by some externally induced forcing $\mathbf{S}(t)$:

$$\mathbf{K}_\sigma \partial_t \sigma = \mathbf{h}(\sigma, \mathbf{S}), \quad (8)$$

where σ is the m -dimensional discretized state vector and $\mathbf{K}_\sigma$ is a diagonal inertia matrix. When the initial conditions are known, the (deterministic) system follows a given trajectory. To help give substance to this general system, we can think of either a global 1D atmospheric model or a single-column GCM, with temperature T and specific humidity Q as the principal components of vector σ and, depending on the model, all the other necessary state-phase variables. Vector $\mathbf{S}$ would then be the incoming solar radiation absorbed by each layer of the reference trajectory.

a. Tangent linear system

Once a reference trajectory ($\mathcal{T}^{\text{ref}}$) is established, small perturbations to system (8) are solutions of the so-called tangent linear system, which is obtained by a simple linearization of Eq. (8) around its trajectory

$$\mathbf{K}_\sigma \partial_t \Delta \sigma(t) = \mathbf{M}_\sigma \Delta \sigma(t) + \Delta \mathbf{F}_x(t). \quad (9)$$

In this equation, $\Delta \sigma(t)$ is the deviation of the perturbed system from the reference trajectory at time t , caused by a perturbation $\Delta \mathbf{F}_x(t)$ starting from $t = 0$, and $\mathbf{M}_\sigma(t) = \partial_\sigma \mathbf{h}(t)$ is the Jacobian matrix of the system. To give examples, the forcing term $\Delta \mathbf{F}_x$ can originate from a change in

input solar radiation or a forcing induced by a change in x , the CO_2 concentration, in which case it becomes $\Delta \mathbf{F}_x = \partial_x \mathbf{h}(x) \Delta x$.

In this article, we consider that the original system is in equilibrium when we apply a small perturbation, which leads to a new equilibrium. In that case, the Jacobian matrix has constant coefficients and a Laplace transformation can be applied to system (9), yielding

$$\mu \mathbf{K}_\sigma \Delta \sigma = \mathbf{M}_\sigma \Delta \sigma(\mu) + \Delta \mathbf{F}_x(\mu), \quad (10)$$

where μ is the Laplace variable and, for the sake of simplicity, the original symbols are kept to denote their Laplace transforms. The advantage of being in the Laplace domain is that the transformed system requires only simple algebraic manipulations to solve partial differential equations.

Such a system does not exhibit any structure in which a feedback loop can be specified. Let us first introduce a test variable, say T_s as the surface temperature, taken to be the last component of vector σ : $T_s = \langle c | \sigma$, where $\langle c |$ is the m th row of the identity matrix.¹ Because we need to clearly identify columns and rows in matrices, we use the Dirac bracket symbols, $\langle x |$ for row and $|x\rangle$ for column; $\langle x | y \rangle$ is a scalar product and $|y\rangle \langle x|$ an external product.

Inserting this structure into Eq. (10) gives

$$\begin{cases} (\mu \mathbf{K}_\sigma - \mathbf{M}_\sigma^b) \Delta \sigma(\mu) = |m_\sigma\rangle \Delta T_s(\mu) + \Delta \mathbf{F}_x(\mu) \\ \Delta T_s(\mu) = \langle c | \Delta \sigma(\mu). \end{cases} \quad (11)$$

This system is mathematically equivalent to the preceding one if we take $\mathbf{M}_\sigma = \mathbf{M}_\sigma^b + |m_\sigma\rangle \langle c|$ —we drop the dependence on μ when it is not necessary. We see that $|m_\sigma\rangle$ might be any part or the rightmost column of $\mathbf{M}_\sigma$. A change in T_s impacts the whole system from the term $|m_\sigma\rangle \Delta T_s$. A feedback mechanism originating from T_s necessarily corresponds to some part of this column matrix.

Let us now focus on the temperature equations. Thanks to the Laplace transformation, this T system can be extracted from the full state-variables system by simple algebraic elimination, as shown in appendix A. Compared to $\mathbf{M}_\sigma$, matrix $\mathbf{M}$ is now reduced in dimension (say $n \times n$) and its constituents are no longer constant as their coefficients are rational fractions in μ : $\mathbf{M}(\mu)$. This reduced system (the T system) represents the conservation of internal energy in each layer, taking all other mechanisms, such as phase change, etc., into account.

¹ $\langle c | = [0, 0, \dots, 1]$.

b. Open-loop system

Because the Planck response plays a very specific role in climatology, we identify, within the global Jacobian matrix $\mathbf{M}$, the LW response part that we represent as matrix $\mathbf{P}$: P_{ij} is the net LW radiation leaving layer i when the temperature of layer j has been increased by 1 K, as in Lu and Cai (2009). The $\mathbf{P}$ and $\mathbf{K}$ are the only constant matrices in this system. Let $\mathbf{M}(\mu) = -[\mathbf{P} - \mathbf{A}(\mu)]$, where matrix $\mathbf{A}$ represents all the other mechanisms in the model. Now, the rightmost column of $\mathbf{A}$, say $|b\rangle$, represents all possible ways a perturbation of T_s can impact the atmosphere. With these notations, the TLS for temperature $\Delta\eta$ reads

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)\Delta\eta = |b\rangle\Delta T_s + \Delta\mathbf{F}_x, \\ \Delta T_s = \langle c|\Delta\eta \end{cases} \quad (12)$$

where $\mathbf{A}^b$ is chosen such that $\mathbf{A} = \mathbf{A}^b + |b\rangle\langle c|$ to ensure the mathematical equivalence with the T system. In the classical presentation, a feedback is associated with a feedback loop that can be opened. Opening the loop means suppressing the effect of the perturbation (here of T_s) on the system. This affects $|b\rangle$, so the open-loop system is obtained by setting $|b\rangle = 0$ in Eq. (12) and keeping the same forcing:

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)\Delta\eta_0 = \Delta\mathbf{F}_x, \\ \Delta\theta_0 = \langle c|\Delta\eta_0, \end{cases} \quad (13)$$

a system that defines the “no-feedback trajectory” whose solution in the surface temperature is given by

$$\Delta\theta_0 = \langle c|(\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)^{-1}\Delta\mathbf{F}_x. \quad (14)$$

c. Feedback-loop response

To restore the simplicity of the feedback gain analysis, we need to get rid of the forcing vector $\Delta\mathbf{F}_x$. This can be done by subtracting Eq. (13) from Eq. (12). We now want to evaluate the deviation $\Delta\eta^r$ of η from the open-loop response: $\Delta\eta = \Delta\eta_0 + \Delta\eta^r$ —where r stands for “return effect” (i.e., what is added by the feedback). The resulting system reads

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)\Delta\eta^r = |b\rangle\Delta T_s, \\ \Delta T_s = \langle c|\Delta\eta^r + \Delta\theta_0 \end{cases} \quad (15)$$

for which the solution in ΔT_s is

$$[1 - \langle c|(\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)^{-1}|b\rangle]\Delta T_s = \Delta\theta_0. \quad (16)$$

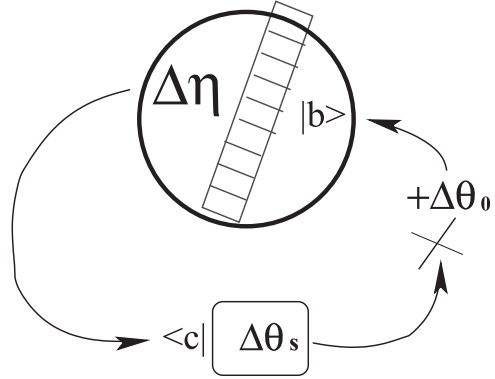


FIG. 3. The feedback-loop-structured system perturbed by $\Delta\theta_0$.

The solution is now identical to Eq. (7) and we may write

$$\Delta T_s = \frac{1}{1-g}\Delta\theta_0, \quad (17)$$

where the feedback gain is given by

$$g = \langle c|(\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)^{-1}|b\rangle. \quad (18)$$

Now, equivalently to system (15), we can also decompose the surface warming into two components: $\Delta T_s = \Delta\theta_s + \Delta\theta_0$, yielding

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)\Delta\eta^r = |b\rangle(\Delta\theta_s + \Delta\theta_0) \\ \Delta\theta_s = \langle c|\Delta\eta^r \end{cases} \quad (19)$$

with its solution

$$\Delta\theta_s = \varrho\Delta\theta_0 \quad \text{with} \quad \varrho(\mu) = \frac{g(\mu)}{1-g(\mu)}. \quad (20)$$

We call ϱ the “effective response,” or the “feedback only” response function. In the Laplace domain, it is the response to the impulse function $\Delta\theta_0$ and is called the “transfer function” in Bode (1945).

Let us comment on these results by following the feedback thread (cf. Fig. 3). Suppose $\Delta\theta_0$ is added to the surface temperature T_s . The troposphere energy budget is perturbed from $|b(\mu)\rangle\Delta\theta_0$ by all the feedback mechanisms represented in the model and the system reacts to redistribute the perturbation among the atmospheric layers: $\Delta\eta(\mu) = [\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b(\mu)]^{-1}|b(\mu)\rangle\Delta\theta_0$. Because of that change in the atmosphere, the initial perturbation to T_s is modified by addition of the term $\langle c|\Delta\eta(\mu)$: the feedback loop is closed. It is characterized by either its gain $g(\mu)$ or its effective-response $\varrho(\mu)$ function.

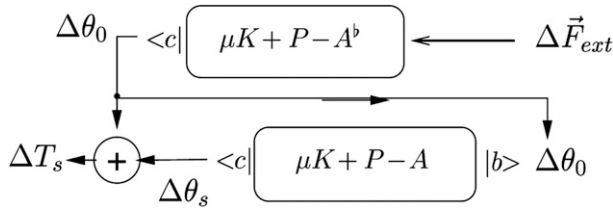


FIG. 4. The forcing system—see Eqs. (14) and (21).

Whereas, in the classical feedback analysis, the perturbation is arbitrarily small, here in a climate sensitivity analysis, it results from the open-loop response to the forcing. It can be pointed out that, with this method, the Planck response to forcing acts as the reference trajectory to the feedback system. Thanks to the linearity hypothesis, the feedback analysis circumvents the difficulty of determining the Planck response.² This is made possible because the three trajectories—reference, Planck, and perturbed—share the same Jacobian matrix.

d. Feedback and forcing

Thanks to the introduction of a basic perturbation, the original forcing vector is replaced by a perturbation in the surface temperature, as shown schematically in Fig. 4.

The figure also shows that $|b\rangle\Delta\theta_0$ plays a similar role to $\Delta\mathbf{F}_x$: it is the specific forcing applied to the surface temperature feedback loop. Eliminating ΔT_s in the first line of system (15) gives

$$(\mu\mathbf{K} + \mathbf{P} - \mathbf{A})\Delta\boldsymbol{\eta}^r = |b\rangle\Delta\theta_0, \quad (21)$$

which shows that the system responds to the forcing vector: $|b\rangle\Delta\theta_0$. Hence, eliminating either ΔT_s or $\Delta\boldsymbol{\eta}^r$ in the same system gives either the forcing response form or the feedback form.

The forcing form [Eqs. (19) and (21)] considers the perturbed system as responding to a specific forcing—specific because of the choice of $|b\rangle$. The feedback form [Eqs. (15) and (16)] focuses on the effect of this mechanism on the system. This equivalence provides a method for determining feedback gains in GCMs.

This introduction to the feedback-loop structure in a model is general and not restricted to the equilibrium

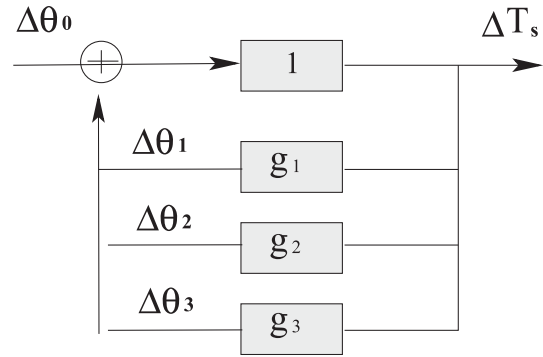


FIG. 5. The classical elementary feedback decomposition (e.g., Peixoto and Oort 1992).

hypothesis; it follows the prescriptions proposed in Lahellec et al. (2008).

4. From exclusive to inclusive feedback

So far, only one feedback loop has been structuring the original system. The application to climate analysis, as we have seen, introduces a decomposition into “elementary feedbacks.” Our questioning of this practice concerns the design of an explicit feedback-loop structure that could associate each elementary response with a specific mechanism.

a. Classical multiple feedback analysis

It is usual to decompose the total feedback gain into separate feedback mechanisms: $g = \sum_i^p g_i$. In Eqs. (15), (16), and (19), $|b\rangle$ represents the way that a surface perturbation impacts the troposphere. If it is possible to identify an elementary mechanism with $|b_i\rangle$ as composing $|b\rangle$, it immediately follows that a partial feedback gain reads

$$g_i = \langle c | (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)^{-1} | b_i \rangle. \quad (22)$$

This defines the classical approach to feedback gains as in Hansen et al. (1984), in which a gain g_i gives the response “that would exist if all other feedbacks were inoperative.” Here, this refers to the open-loop model with the Jacobian matrix $(\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)$. This approach is exemplified by Fig. 5, which shows the usual feedback loops in parallel: we have added the decomposition $\Delta T_s = \Delta\theta_0 + \sum_i \Delta\theta_i$ to this diagram, because it is defined in relation with the gains g_i as $\Delta\theta_i = g_i \Delta T_s$.

From its beginning, the feedback approach to climate analysis has attempted to decompose the global temperature response into components. This was first done by Hansen et al. (1984) and formally detailed by

² Stephens (2005): “Since we cannot generally observe the real climate system with feedbacks turned off, any use of observations for this purpose requires assumptions that are generally hard to justify a priori.”

Schlesinger (1986). An application is found in Hansen et al. (1984) but we will consider a possible confusion in practical determinations. Also, LeTreut et al. (1994), in their analysis of a global warming experiment, found it “informative to split the surface warming” into components, with a method we will show appropriate. We have found no other consistent application before Dufresne and Bony’s (2008) comparison study of 12 models participating in CMIP3–Fourth Assessment Report (AR4) in support of the IPCC missions. In parallel, the CFRAM approach of Lu and Cai (2009) justified this decomposition from an independent approach.

With the structure in Fig. 5, each temperature warming component is associated with a feedback mechanism because each gain is associated with a column matrix $|b_i\rangle$. However, as we will see, no loop can be opened that would result in a shift $\Delta\theta_i$ of ΔT_s . In contrast, noting that the global effective response given by Eqs. (18) and (20) can also be rewritten, as shown in Eqs. (B1) and (B2), as

$$q = \langle c | (\mu \mathbf{K} + \mathbf{P} - \mathbf{A})^{-1} | b \rangle, \quad (23)$$

where the full Jacobian matrix now appears, allows the decomposition

$$q_i(\mu) = \langle c | [\mu \mathbf{K} + \mathbf{P} - \mathbf{A}(\mu)]^{-1} | b_i(\mu) \rangle. \quad (24)$$

We call it the “response form” of q , which explicitly shows the additivity of the effective responses. The associated feedback form hence becomes

$$\begin{aligned} q_i(\mu) &= \frac{\langle c | [\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^b(\mu)]^{-1} | b_i(\mu) \rangle}{1 - \langle c | [\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^b(\mu)]^{-1} | b(\mu) \rangle} \\ &= \frac{g_i}{1 - g} = \frac{g_i}{1 - \sum_j g_j}, \end{aligned} \quad (25)$$

which refers the perturbation components to the basic perturbation: $\Delta\theta_i = q_i \Delta\theta_0$. Notice that Lu and Cai (2009) introduce what they call an “additive gain” $\tilde{g}_i$, which can be shown to be our effective response q_i at the steady state— $\mu = 0$ in Eq. (25).

The last form of the effective response in (25) clearly shows that an elementary gain g_i cannot be set to zero while keeping the additivity of the perturbation components—the denominator will also change. This problem explains the difficulty of associating $\Delta\theta_i$ with a feedback-loop structure where the opening of the loop would shift the response by $\Delta\theta_i$. Another approach is needed to justify the decomposition, and this is the object

of our “inclusive” approach, which, if we are to be consistent, qualifies the classical one as “exclusive.”

b. Inclusive feedback components

Up to now, the perturbation has acted as a tool exciting one mechanism at a time to evaluate its importance in the truncated system response—truncated because the matrix $\mathbf{A}^b + |b_i\rangle\langle c|$ does not represent the full feedback dynamics [see Eq. (22)]. In climate sensitivity studies, the situation is different: the question concerns the effect of some mechanism (i.e., its influence within the realistic global ΔT_s response). In other words, we are looking to associate one of the $|b_i\rangle$ with a feedback loop or, to put it differently, we want the full system to respond to the specific forcing $|b_i\rangle\Delta\theta_0$.

The algebraic structure corresponding to this forcing is now, instead of Eq. (19),

$$\begin{cases} (\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^i) \Delta \boldsymbol{\eta}_i^r = |b_i\rangle (\Delta\theta_i + \Delta\theta_0) \\ \Delta\theta_i = \langle c | \Delta \boldsymbol{\eta}_i^r \end{cases}, \quad (26)$$

where $\mathbf{A}^i = \mathbf{A} - |b_i\rangle\langle c|$ represents the reaction of the system when only the i th loop is opened. We obtain a new forcing structure replacing that of Eq. (11). Eliminating $\Delta \boldsymbol{\eta}_i^r$, the solution reads

$$\begin{aligned} (1 - g_i^*) \Delta\theta_i &= g_i^* \Delta\theta_0, \quad \text{with} \\ g_i^* &= \langle c | (\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^i)^{-1} | b_i \rangle. \end{aligned} \quad (27)$$

Now, one mechanism is associated with the feedback loop within the full system, and we have a temperature-change component associated with this feedback:

$$\Delta\theta_i = \frac{g_i^*}{1 - g_i^*} \Delta\theta_0. \quad (28)$$

The new feedback structure corresponds to the right panel of Fig. 6. Note that $\Delta\theta_0$ in Eq. (26) is replaced by $\Delta\theta^i = (g - g_i)/(1 - g + g_i)$, which is the open-loop response (as $\Delta\theta^2$ in the figure) of the inclusive feedback loop.

In this structure, a feedback loop is associated with one mechanism and can be opened, but now the perturbation seen by this loop is primarily altered by the response of the rest of the system to the original perturbation. Also, it is necessary to consider as many loop-structured systems as there are p mechanisms listed. Each feedback loop is characterized unambiguously by the effective response q_i and its corresponding inclusive gain g_i^* . It depends neither on the Planck response nor on whether other feedbacks are elicited or not: they are

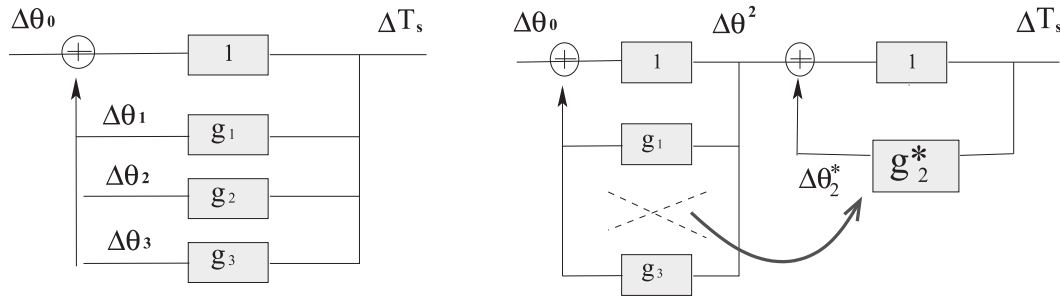


FIG. 6. Process 2 analyzed (left) exclusively and (right) inclusively.

intrinsic parameters characterizing a feedback mechanism in the dynamical system.

Unlike the exclusive feedback gains, the inclusive gains do not add up: it is the effective responses that do. It is worth pointing out that the gain g_i^* does not give the response when all other mechanisms are frozen; it gives the response associated with one mechanism, all others being active. This is why we call this feedback gain inclusive and accordingly give the name exclusive to the gain defined in text books (e.g., Peixoto and Oort 1992) and adopted by Hansen et al. (1984) in which, to our mind, a confusion between the two definitions has been made in their temperature decomposition.

Compared to the inclusive feedback gain in Eq. (27), the exclusive feedback gain is $g_i = \langle c | (\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^b)^{-1} | b_i \rangle$, as $\mathbf{A}^b$ represents the response of the system when all feedback mechanisms are frozen. So, the presence of part of the $\mathbf{A}$ matrix discriminates between inclusive and exclusive gains.

c. Link between exclusive and inclusive analyses

When $|b\rangle = \sum_{i=1}^p |b_i\rangle$ is the rightmost column of $\mathbf{A}$, with a total gain $g = \sum_{i=1}^p g_i$, the relations between the two modes are easily found as

$$g_i^*(\mu) = \frac{g_i(\mu)}{1 - g(\mu) + g_i(\mu)}$$

$$q_i(\mu) = \frac{g_i(\mu)}{1 - g(\mu)} = \frac{g_i^*(\mu)}{1 - g_i^*(\mu)}; \quad \text{also:}$$

$$\Delta\theta_i(\mu) = \frac{g_i^*(\mu)}{1 - g_i^*(\mu)} \Delta\theta_0(\mu) = \frac{g_i(\mu)}{1 - g(\mu)} \Delta\theta_0(\mu), \quad (29)$$

where the dependence in μ has been recalled. See also Fig. 6. One can verify that $\Delta\theta_i = g_i \Delta T_s$, which are the inclusive feedback components (IFCs) of the total temperature change ΔT_s .

Schlesinger (1986) gave the preceding expressions of the gain and the IFCs at equilibrium, although with no application. Perhaps his choice of using classical electrical

engineering terminology, which has been abandoned in the climate community, could explain why his suggestions have not been followed. We have only extended his results to the dynamics of transition to the new equilibrium but it has to be emphasized that, while the equilibrium response is trivial to compute, the dynamics concealed in the preceding formulas established in the Laplace domain needs an inversion to the real domain, which is not trivial. It may be worked out using the method described in Hallegatte et al. (2006) for instance.

Two practices, corresponding to the exclusive and the inclusive approaches to feedback analysis, can be found in the literature. We take the sea ice albedo feedback as an example because the two approaches give very different results in this case. Many methods use the radiation code only to assess the change (PRP methods in general) in SW TOA budget when the sea ice extent has changed during a $[\text{CO}_2]$ -doubling scenario, the practical application varying by the way time and spatial averages are taken. In such a strategy, no effect of the change in albedo is found in the bands of latitude unaffected by sea ice. The other mechanisms do not respond to the perturbation of the albedo and the method pertains to the exclusive approach—see Wetherald and Manabe (1988) for instance.

In contrast, the “feedback suppression” methods compare two runs where the sea ice extent is maintained at the reference value or free to evolve: the albedo feedback loop is open or closed within the global system. In this case, all regions are concerned because of the redistribution of SW budgets in the troposphere impacting the sea-iced zones in the vertical, and consequently, the global atmosphere dynamics. This is an inclusive approach; see for instance Hall (2004) and Ingram et al. (1989). Albedo feedback factors can vary by a factor of 2 between the two methods (Ingram et al. 1989). Such uncertainty on the albedo feedback propagates to the other feedbacks because the sum of all feedback factors is known independently.

The determination of an IFC from the suppression method is a source of difficulty, encountered by Hall and

Manabe (1999) in their analysis of the water-vapor (WV) feedback. By opening the WV feedback loop, one might expect to obtain a response having a direct relation with the corresponding IFC, but this is not the case. Opening the loop gives indeed $\Delta\theta^i$, the open-loop response as in Fig. 6, which enters in the relation with the IFC as³

$$\Delta\theta_i = \left(\frac{\Delta T_s}{\Delta\theta^i} - 1 \right) \Delta\theta_0. \quad (30)$$

Taking the values from Ingram et al. (1989), with $\Delta T_s = 5.17$ K, and $\Delta\theta^i = 4.19$ K as the open-loop response, with the standard value of $\Delta\theta_0 = 1.2$ K, this gives a sea ice albedo temperature component $\Delta\theta_\alpha$ of 0.28 K, close to the mean CMIP3 value of 0.3 K in Dufresne and Bony (2008) for the surface albedo feedback. Opening the loop in Ingram et al. (1989) gives $\Delta T_s - \Delta\theta_\alpha = 1$ K instead. It would correspond to $\Delta\theta_2^*$ in Fig. 6. Equation (30) shows that it may be nontrivial to disentangle an IFC from the interaction between feedback and forcing when the feedback loop is not explicit. Also, as far as it is possible to check in the article, Hansen et al. (1984) determined their warming components with only one active feedback at a time, leading to an apparent discrepancy between their Fig. 6 and Table 1. Notice that in contrast, CFRAM intimately associates these components with mechanisms.

5. From feedback views to climate sensitivity

We now derive an explicit form of the approach to climate sensitivity as initiated by Wetherald and Manabe (1988), leading to the PRP techniques.

a. Linking climate sensitivity

To link TOA fluxes and surface temperature, we decompose the system, extracting the surface temperature equation. This leads us to introduce the submatrix notation:

$$\mathbf{P} = \begin{bmatrix} \bar{\mathbf{P}}, & |\bar{p}\rangle \\ \langle\bar{p}|, & \bar{p} \end{bmatrix}; \quad \mathbf{A} = \begin{bmatrix} \bar{\mathbf{A}}, & |\bar{a}\rangle \\ \langle\bar{a}|, & \bar{a} \end{bmatrix}, \quad (31)$$

where the last row and rightmost column related to the surface temperature are now identified in the Jacobian matrices, and hence in the two submatrices of $\mathbf{A}$:

$$\mathbf{A}^s = \begin{bmatrix} \bar{\mathbf{A}}, & |0\rangle \\ \langle\bar{a}|, & 0 \end{bmatrix}, \quad |b\rangle = \begin{bmatrix} |\bar{a}\rangle \\ \bar{a} \end{bmatrix} \quad (32)$$

and their decomposition; for example, $|\bar{a}\rangle = \sum_i |\bar{a}_i\rangle$, etc. With this notation, as shown in Eq. (C11), the new form of the feedback gain reads

$$g_i(\mu) = \frac{\bar{a}_i - \langle\bar{p} - \bar{a}|(\mu\bar{\mathbf{K}} + \bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{a}_i\rangle}{\mu\bar{k} + \bar{p} - \langle\bar{p} - \bar{a}|(\mu\bar{\mathbf{K}} + \bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{p}\rangle}, \quad (33)$$

with the components of $\mathbf{A}$ also depending on μ , which at equilibrium reads

$$g_i = -\frac{\bar{a}_i - \langle\bar{p} - \bar{a}|(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{a}_i\rangle}{\bar{p} - \langle\bar{p} - \bar{a}|(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{p}\rangle}, \quad (34)$$

which can be compared to Eq. (22). What can be learned from this new expression of the gain?

- The denominator is the no-feedback net surface budget—see Eq. (C6):

$$-\lambda_0 = \bar{p} - \langle\bar{p} - \bar{a}|(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{p}\rangle. \quad (35)$$

A perturbation of the surface temperature causes a radiative emission $\bar{p}$ from the surface layer, and an absorbed part ($-\langle\bar{p}\rangle$) in the tropospheric layers. This absorbed radiation induces a tropospheric temperature variation $\Delta\bar{\eta} = -(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{p}\rangle$. In turn, the surface layer receives $-\langle\bar{p} - \bar{a}|(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{p}\rangle$ net budget. At the tropospheric equilibrium, this equals the outgoing TOA budget—with our sign convention, λ_0 is negative.

- Also, the numerator is the net perturbed budget of both the surface layer and TOA for the i th feedback mechanism—incoming budget (i.e., λ_i)

$$\lambda_i = \bar{a}_i - \langle\bar{p} - \bar{a}|(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}|\bar{a}_i\rangle; \quad (36)$$

see Eq. (C7) for λ_T .

- In either case, the surface perturbation is propagated to the troposphere through $|p\rangle$ or $|b_i\rangle$, and the response involves $(\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1}$, the troposphere with all its internal processes, also acting in the feedback to surface via $\langle\bar{p} - \bar{a}|$.

Note that λ_0 defines a novel Planck response $\Delta R = -\lambda_0\Delta\theta_0$ [see Eqs. (C5) and (C6)] that we name the “surface Planck” response ΔR and sensitivity factor λ_0 . It is the TOA response to a pure radiative forcing emanating from the surface. The Planck response will be numerically determined in section 6.

Finally, the new form of the gain at equilibrium is found as $g_i = -\lambda_i/\lambda_0$, justifying the way sensitivities and feedback gains are linked in Eq. (4). This identity of

³ See detailed derivation in the supplemental material.

results from the two theories is made possible because the PRP focuses on the very core of the greenhouse effect: a change in the contribution to the TOA budgets. This becomes more evident when we see that Slingo et al. (2000)'s normalized greenhouse effect $\mathcal{N}$, which points to the part of the surface emission trapped in the troposphere, here reads

$$\mathcal{N} = \frac{\langle \bar{p} - \bar{a} | (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} | \bar{p} \rangle}{\bar{p}}; \quad (37)$$

so, at equilibrium, this part trapped in the troposphere is also the energy fed back to the surface by the perturbed system.

The link between the inclusive and exclusive analysis may be summed up in⁴

$$\begin{aligned} \text{feedback: } g_i &= \frac{\lambda_i}{\lambda_0} \quad g_i^* = -\frac{\lambda_i}{\lambda_T - \lambda_i} \\ \text{IFC: } \Delta\theta_i &= -\frac{\lambda_i}{\lambda_0} \Delta T_s \quad \Delta\theta_i = -\frac{\lambda_i}{\lambda_T} \Delta\theta_0. \end{aligned}$$

We recall that $\Delta T_s = \sum_i \Delta\theta_i + \Delta\theta_0$. It is essentially the inclusive expression of the IFC $\Delta\theta_i$ that LeTreut et al. (1994) used to split their global warming results into components.

It should nevertheless be noted that this identity of results between the PRP analysis and the feedback gain approach is only valid at equilibrium. In other situations, the surface layer divergence does not equal the outgoing TOA budget. Care should thus be taken when dealing with nonsteady analysis and we will come back to this difficulty in the next paragraph.

b. From PRP to the regression method

Extending the PRP approach, Gregory et al. (2004) proposed an innovative method to determine the global climate sensitivity to forcing from regression between the amplitudes of time series in $N(t)$ against $\Delta T_s(t)$, putting into application Slingo et al. (2000)'s program: “a more direct measure of the feedback should come from the interannual variability of global-mean quantities, because this timescale and space scale is more appropriate for such a global phenomenon.” For long-enough-lasting forcing, they claim that the linearity hypothesis allows the relation

$$-N = \lambda_T \Delta T_s + \text{forcing} \quad (38)$$

to come into play. In these conditions, the slope of the asymptotic regression curve gives the global sensitivity factor λ_T and its y intercept is the constant forcing term. The method proved to be acceptable for global annual-mean results from 20 atmosphere–ocean general circulation models (AOGCMs) involved in CMIP3 (Forster and Taylor 2006). This method is particularly interesting because it introduces the possibility of using data from the generation of satellites launched in the 2000s. Forster and Gregory (2006) gave promising preliminary results with 6 years of satellite data and estimates of the climatic forcings. We shall see, however, that this method can lead to severe difficulties.

We want to examine under what conditions this technique can bring objective information on climate sensitivity. Starting from Eq. (12) and applying the submatrix decomposition [Eq. (31)], a Taylor development in μ gives⁵ for the long-lasting constant forcing ($\mu \rightarrow 0$):

$$\begin{aligned} N &= [\bar{p} - \bar{a} - \langle \bar{p} - \bar{a} | (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} | \bar{p} - \bar{a} \rangle] \Delta T_s \\ &\quad - [\Delta \bar{f}_x - \langle \bar{p} - \bar{a} | (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} \Delta \bar{\mathbf{F}}_x] + \mathbf{O}(\mu). \end{aligned} \quad (39)$$

This expression validates the regression method, because

$$-\lambda_T = \bar{p} - \bar{a} - \langle \bar{p} - \bar{a} | (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} | \bar{p} - \bar{a} \rangle \quad (40)$$

[cf. Eq. (C7)], and the surface forcing term

$$f_0 = [\Delta \bar{f}_x - \langle \bar{p} - \bar{a} | (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} \Delta \bar{\mathbf{F}}_x] \quad (41)$$

is the impact of the vector forcing on the surface temperature⁶ $[\Delta \mathbf{F}_x = (\Delta \bar{\mathbf{F}}_x, \Delta \bar{f}_x)^\dagger]$. However, numerical investigation of the regression method with the 1D model Climate Simplifié (ClimSI)—as in Hallegatte et al. (2006)—shows that the asymptotic straight line is only reached after a few years, because the fast feedbacks, such as the fast pole of the negative WV feedback found in the cited article, still intervene significantly along with the slow pole after even the first year of application of the perturbation. Such an interfering effect is also reported in Forster and Taylor (2006).

Conditions to reach the asymptotic straight line can be deduced from a Lyapunov analysis of fraction:

$$\lim_{\mu \rightarrow 0} \frac{N(\mu)}{\Delta T_s(\mu)} = \frac{\langle 1 | (\mathbf{P} - \mathbf{A})(\mu \mathbf{I} + \mathbf{P} - \mathbf{A})^{-1} \Delta \mathbf{F}_x \rangle}{\langle c | (\mu \mathbf{I} + \mathbf{P} - \mathbf{A})^{-1} \Delta \mathbf{F}_x \rangle}. \quad (42)$$

With the eigen elements of matrix $\mathbf{P} - \mathbf{A} = E[w]_d F^\dagger$, cross norm $F^\dagger E = \mathbf{I}$ between the left and right

⁴Details in appendix C on the decomposed feedback system.

⁵Full development in the supplemental material.

⁶Compare to Eq. (36) in the supplemental material.

eigenvectors, and $[w]_d$ the diagonal eigenvalues matrix, one finds for the impulse forcing that

$$\begin{cases} N(t) = \sum_k w_k e^{w_k t} \langle 1 | e_k \rangle \langle f_k | \Delta \mathbf{F}_x \rangle \\ \Delta T_s(t) = \sum_k e^{w_k t} \langle c | e_k \rangle \langle f_k | \Delta \mathbf{F}_x \rangle \end{cases} \quad (43)$$

Each function in the real domain is time dependent through a weighted sum of exponentials—possibly as factors of sine functions for intrinsic oscillatory mechanisms. For the constant slope to be reached, both functions must depend on the same function of time, which, in general, only occurs when all but the last exponential have died away.⁷ The constant forcing responses are the integrals of the previous ones, adding a constant that will be reached asymptotically.

The conclusion is that the straight line is reached as soon as the leading Lyapunov exponent is left alone. This can be checked by looking for the exponential decay of ΔR and ΔT_s with time, hence allowing that exponent to be determined. So, in these conditions, while matrices in Eq. (39) are constant, ΔT_s still has to evolve for a regression to be useful. The evolution of ΔT_s is then given by

$$\bar{k} d_t \Delta T_s = -N = \lambda_T \Delta T_s + f_0 \quad (\text{L2Obs}). \quad (44)$$

We call Eq. (44) the L2Obs relation: in addition to the linear hypothesis, it relates the two observables N and ΔT_s over long time scales. We can conclude that the regression method is applicable for systems where the thermal capacitance ($\bar{k}$) attached to the test variable is much greater than the others cumulated. In these conditions, the atmosphere is quasi-stationary drifting with the surface temperature. This equation further shows that the leading Lyapunov exponent corresponds to the e -folding time

$$\tau = -\bar{k}/\lambda_T. \quad (45)$$

Such property should allow the apparent heat capacitance to be retrieved, and applications are found in Knutti et al. (2008), Schwartz (2007), and in Part II of the present paper. But again, precautions have to be taken to check for coherence with the time evolution of the variables. In the real context as well as in the GCM analyses, noisy regressions can lead to biases, as argued in Spencer and Braswell (2010), and regressing on short-term variations may lead to improper analyses, such as

the misdiagnosis on “the short term” cloud radiative forcing feedback in Dessler (2010).

6. On the regression technique

The regression technique is now currently used in the CMIP analyses for comparing global feedback factors between GCMs. It is also linked to “pattern-scaling analyses” (i.e., Mitchell et al. 1999) as characterizing the dynamical structures of the climate warming. In this context, our Lyapunov argument on the validity of regression techniques needs an illustration to make it clear. This is an opportunity to show how in principle a TLS could be implemented in GCMs.

a. Regression on the Planck response

Because most GCMs are classically using a double calling of the radiation code to determine a given forcing, one has just to integrate the difference in radiation divergence obtained to get the TLS. This is easy to obtain: each time step along the run, the radiation code is called with all the arguments needed to compute the divergence of radiation energy at each level: surface and TOA. In the Laboratoire de Météorologie Dynamique Zoom GCM (LMDZ), the Morcrette code is used (Morcrette 1991), which gives two vectors $\mathbf{h}$ and $\mathbf{c}$ (heat and cold) as output, representing the temperature trend (K day^{-1}) of heating by solar flux and cooling by LW radiation, respectively. Atmospheric temperatures are then advanced from time k to $k + 1$ for the time step δt :

$$\mathbf{T}_{k+1} = \mathbf{T}_k + [\mathbf{h}(x, \mathbf{T}) - \mathbf{c}(x, \mathbf{T})] \delta t. \quad (46)$$

Among all of the arguments, we have explicitly written x as the CO_2 concentration. By definition, the true Planck response is the variation of temperature when $\mathbf{h} - \mathbf{c}$ changes with $\mathbf{x}$. The corresponding local TLS is given by differentiating Eq. (46):

$$\Delta \mathbf{T}_{k+1} = \Delta \mathbf{T}_k + [(\partial_x \mathbf{h} - \partial_x \mathbf{c}) \Delta x + (\partial_T \mathbf{h} - \partial_T \mathbf{c}) \Delta \mathbf{T}] \delta t, \quad (47)$$

where two Jacobian matrices of the radiation code appear, one of them being matrix $\mathbf{P} = -(\partial_T \mathbf{h} - \partial_T \mathbf{c})$. The equivalent formal system can be written

$$\begin{cases} \partial_t \mathbf{T} = f(\mathbf{T}, x) \\ \partial_t (\mathbf{T} + \Delta \mathbf{T}) = f(\mathbf{T} + \Delta \mathbf{T}, x + \Delta x) \end{cases} \quad (48)$$

and the TLS is obtained by first-order development and subtraction

$$\begin{aligned} \partial_t (\mathbf{T} + \Delta \mathbf{T}) &= f(\mathbf{T}, x) - \mathbf{P} \Delta \mathbf{T} + \partial_x f \Delta x \\ \partial_t \Delta \mathbf{T} &= -\mathbf{P} \Delta \mathbf{T} + \partial_x f \Delta x \end{aligned}, \quad (49)$$

⁷ Ibidem.

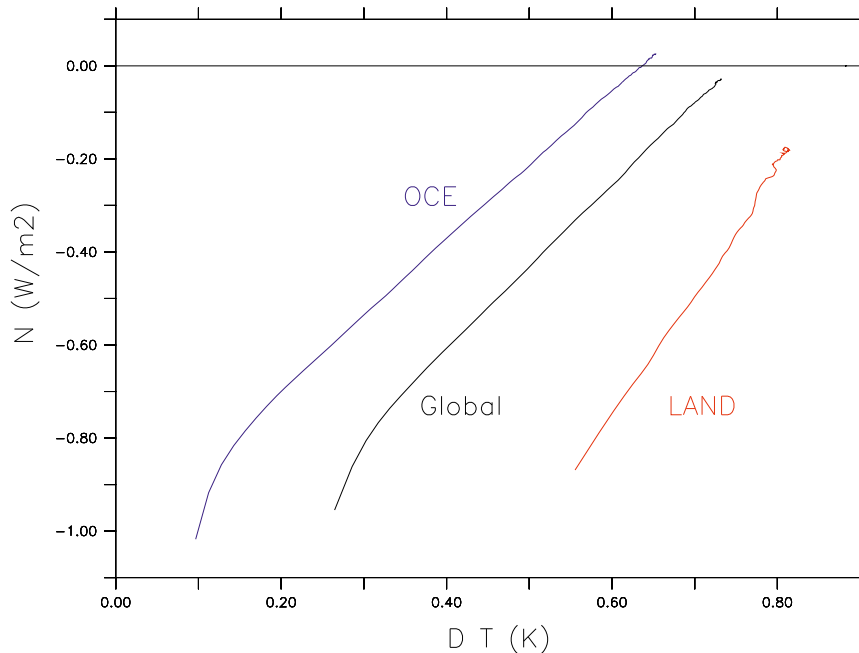


FIG. 7. Regression on the Planck response—12-month low-pass filtered.

where the forcing term is $\partial_x \mathbf{f} \Delta x = \Delta(\mathbf{h} - \mathbf{c})$. At equilibrium, this system is identical to the corresponding Eq. (7) of Cai and Lu (2009). The finite difference version reads

$$\Delta \mathbf{T}_{k+1} = \Delta \mathbf{T}_k + \{[\mathbf{h}(2x, \mathbf{T} + \Delta \mathbf{T}) - \mathbf{c}(2x, \mathbf{T} + \Delta \mathbf{T})] - [\mathbf{h}(x, \mathbf{T}) - \mathbf{c}(x, \mathbf{T})]\} \delta t. \quad (50)$$

The last difference on the rhs is the response of the radiation code along the trajectory; the other gives the response to the doubling of $[\text{CO}_2]$ and requires a second call to the radiation code. The difference in brackets is nowadays systematically used to evaluate forcings, and we just added time integration to obtain the temperature response. It is the true Planck response because no other feedback is included—other than the Planck feedback, which corresponds to the presence of ΔT in Eq. (50) and ensures numerical stability. We have not written the corresponding surface temperature advance as, obviously, it is also computed with the difference of surface divergences and a possibly different time scheme. This procedure gives the possibility to determine the Planck response dynamically and represents the most simple TLS that can be implemented in GCMs—in fact, the $\mathbf{P}$ matrix in Eq. (49) is varying in time, which we call circulating this TLS (CTLs).⁸ Note that matrix $\mathbf{A}^b$ in

Eq. (13) is no longer necessary because we determine here the full (3D) Planck response, not only a surface-Planck response. We use this linear model to illustrate our analytical results concerning the regression method.

Figure 7 gives the results of the implementation of the procedure in LMDZ5A (cf. Hourdin et al. 2013) run with an oceanic slab model. The global response (in black) to the abrupt $2\times\text{CO}_2$ forcing is slowly evolving because of the oceanic surface inertia (50 m in this example).

For the L2Obs relation to be valid, we have to check that the straight line corresponds to exponential functions of $N(t)$ and $\Delta T_s(t)$. Figure 8 shows that both parameters are following the integral-exponential function $a + b[1 - \exp(-t/\tau)]$ after a few months of transition, with an e -folding time close to 36 months. This allows us to validate the L2Obs relation as giving $\lambda_0 = -2 \text{ W m}^{-2} \text{ K}$ from the slope. This value differs from the classical $-3.2 \text{ W m}^{-2} \text{ K}$ in Dufresne and Bony (2008). The corresponding Planck surface warming is of 0.75 K instead of 1.2 K.

This important discordance shows the radical difference of our determination of the Planck response with PRP. In PRP, using the “stratosphere relaxation” after the abrupt forcing in GCMs leads to us incorporate all other perturbed fast mechanisms in the response as precipitation, clouds, etc. By construction, no other mechanism is perturbed with the CTLs approach—the reason we call it the “true Planck response.” Another difference with PRP is that the regression technique

⁸ Compare to the supplemental material.

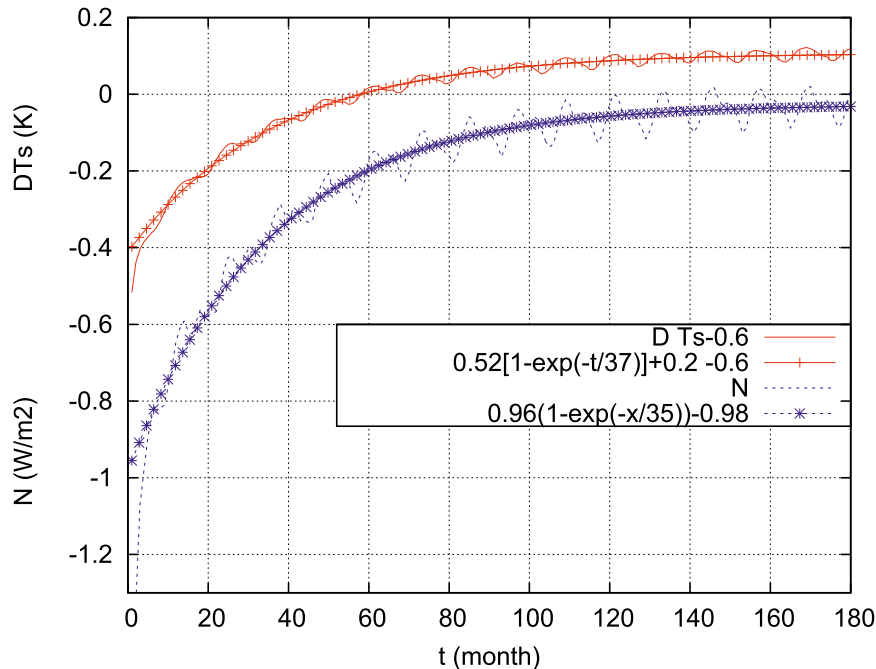


FIG. 8. Test of exponential evolution of N and $\Delta\theta_s$ —monthly values and fit.

applied to the Planck response allows us to make the distinction between the long-term and the fast response. In our opinion, the long-term forcing should be the one used as the gauge to evaluate the other long-term feedback gains, leaving the fast response to another analysis framework.

b. Using the Lyapunov criterion

To unveil the meaning of Eq. (42), we trace the Lyapunov eigenvector by plotting $|y_0\rangle$, with coefficients $\Delta\eta_i/\Delta T_s$, for a few levels in Fig. 9.

The asymptotic constancy for the yearly values is attained within a few years for the troposphere temperature but takes more than 5 years in the stratosphere. The leading Lyapunov criterion thus strongly constrains the regression method and, depending on the desired precision, excludes any attempt to link fast features with climate warming characteristics using that method, because fast features are incompatible with the constancy of $|y_0\rangle$.

We also claimed that the checking for N and $\Delta T_s(t)$ having the exponential behavior allows us to retrieve the apparent global surface inertia. Following Eq. (45), the e -folding time of 36 months found from fit is smaller than the value given by $-\bar{k}/\lambda_0$ for 50 m of seawater, which gives 40 months. So surfaces other than oceanic contribute to lower the surface inertia. On the global scale, however, this 36-month time scale is long enough for the global atmosphere to appear slaved to the global surface

warming. Figure 7 also gives the contribution of the different surfaces to the global regression. On the continental surface, the slope is quite different—about $-3 \text{ W m}^{-2} \text{ K}$. The spatial dispersion of the slope is shown in Fig. 10 for values inferior to $10 \text{ W m}^{-2} \text{ K}$, which corresponds to the surface inertia of 15 m of water—about the limit to the validity of the L2Obs relation.

Large parts of continental and high-latitude regions are off this limit (white grid points), thus requiring distinctive pattern-scaling analyses. Our conclusion is that the regression technique is valid on the global scale because of the predominance of oceans. It could fail, for instance, to characterize the long-term climate sensitivity during glacial cycles of the planet.

7. Concluding discussion

The first part of this paper was devoted to questioning the global warming analysis procedures in common use in the CMIP community. We built a compact formalism that was sufficiently explicit to review current methods for the determination of climate sensitivity to forcing. Although we have not formally introduced the space and time averaging procedure in this article, the tangent linear system (TLS) obtained might be thought of as representing the conceptual model underlying global-mean analyses.

Using this formal framework allowed the feedback approach and the PRP sensitivity analysis to be linked

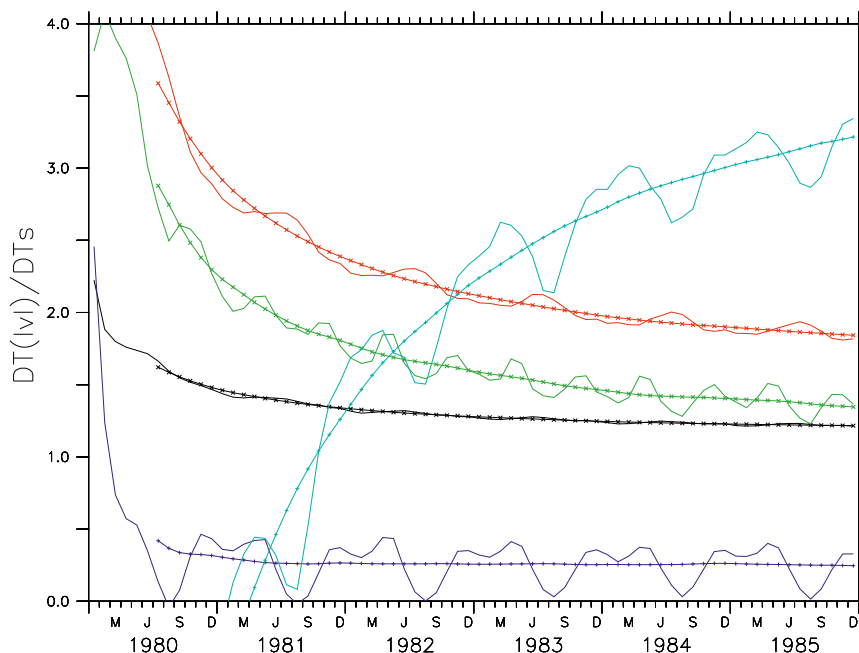


FIG. 9. Five atmospheric levels l of ratio $\Delta\eta/\Delta\theta_s(t)$ with global and monthly–12-month low-path filtered values: 1000 (black), 850 (red), 140 (green), 70 (blue), 30 (magenta) hPa.

rigorously but with two limitations: we first had to introduce a specific forcing as the surface Planck forcing and we also had to admit that we could only represent mechanisms that were directly perturbed from the surface. With these limitations, however, feedback and sensitivity analyses are formally equivalent, not only at equilibrium but also along the perturbed trajectory. We have shown that the decomposition into partial feedbacks leads to two modes of analysis—the exclusive one in the PRP approach and the inclusive one in the suppression methods—in their intention. To our knowledge, only CFRAM can be interpreted as a correct inclusive approach—we discuss this in detail in Part II. We have also demonstrated that the surface temperature change can be decomposed into what we call inclusive feedback components (IFCs). Using different model structures, each sensitivity factor can be associated with a feedback loop and an IFC, without loss of physical consistency. We emphasized the need to explicit the feedback loop involved in a feedback suppression method to properly interpret the results.

To be effective, however, this formal equivalence between the inclusive and exclusive approaches would need a conceptual model able to associate each perturbed field in the classical decomposition—temperature, water vapor, clouds, and albedo—with specific mechanisms initiated from the surface, and here lies the real difficulty. Nevertheless, this does not prevent

the PRP decomposition from being used as the simplest procedure to make comparisons between GCMs and the second method for investigating the role of climate mechanisms in a more physical analysis.

In a second part of this paper, we made use of the dynamical aspects of our formal model to show that the regression method is a very pertinent tool for global analyses, as long as specific criteria are verified. These criteria are quite constraining, leading to time-scale filtering of the global climate evolution over at least 5 yr, as found with the numerical illustrations in section 6. Yet, at the same time, these constraints palliate the limitations of the strict formal approach and extend the range of mechanisms that can be analyzed in GCMs. The fact that, at the Lyapunov horizon, the global atmosphere appears slaved to surface temperature suggests that a formal extension of our strict analysis is possible, as will be shown in Part II. This allowed us to introduce a new definition of the true Planck response with a method featuring a numerical framework that could be eventually applied to other feedbacks. In this paper, it served as a support to illustrate our abstract arguments based on a Lyapunov analysis of the regression technique. Adding the dynamical analysis to the regression method as we did has a double advantage: it checks that the Lyapunov constraint is attained and also

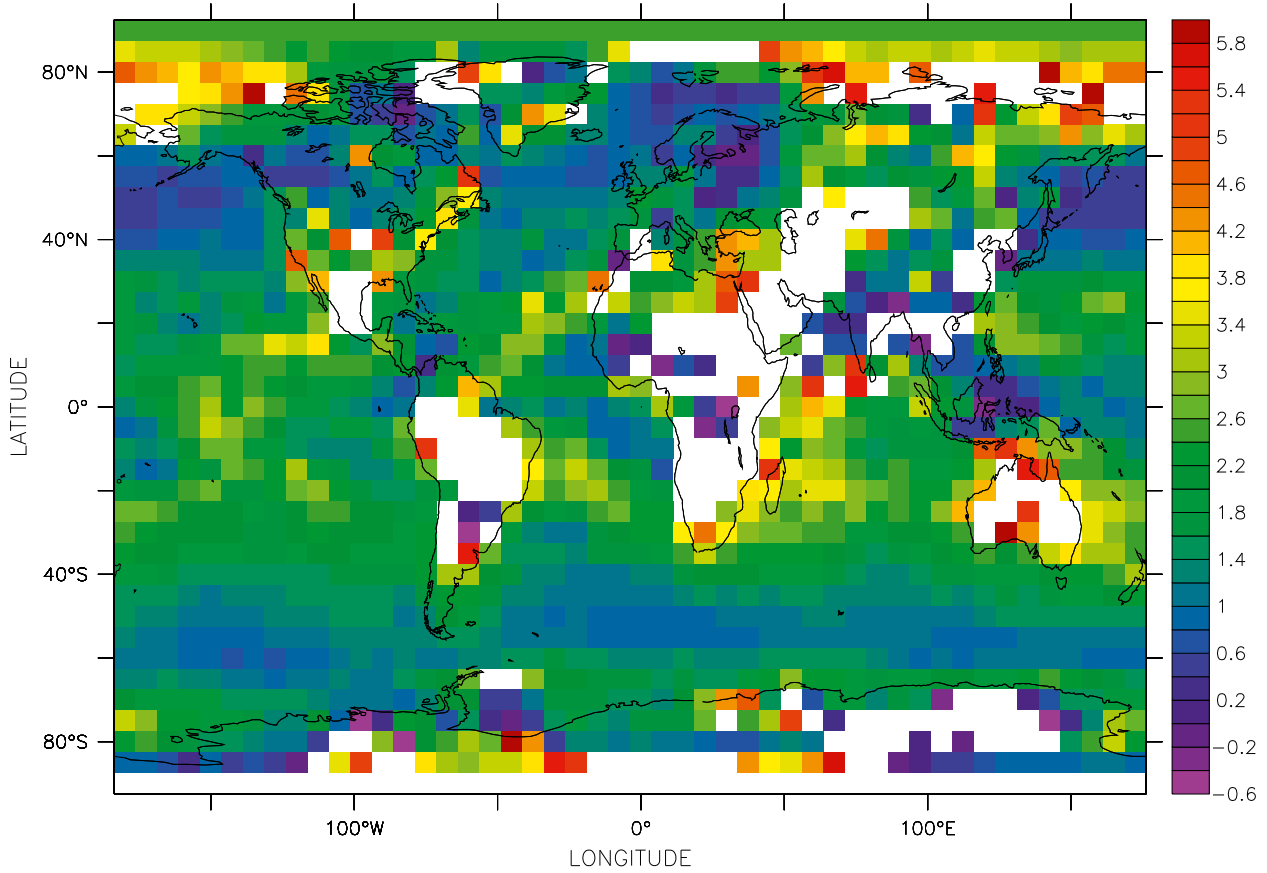


FIG. 10. Dispersion of the slope given by the L2Obs relation on the globe ($\text{W m}^{-2} \text{K}$).

determines the ocean heat capacity in a coherent way. This is particularly relevant for CMIP5 analyses as we show in Part II.

Acknowledgments. The authors gratefully acknowledge useful discussions with J. Y. Grandpeix and S. Hallegatte. Many thanks to O. Talagrand for his support. We wish to thank Susan Becker for careful proofreading of the English version. We are particularly grateful for detailed corrections and suggestions from our editor Ming Cai and also from an anonymous reviewer.

APPENDIX A

Extending the T System

Let the vector ζ represent all state variables, other than the temperatures η , defining the system of Eq. (11)—pressure, wind field, water content, etc. The feedback system reads

$$\begin{cases} \mu \Delta \zeta = A_q \Delta \zeta + B_{qt} \Delta \eta + |b_{qt}\rangle (\Delta \theta_s + \varpi) \\ (\mu K + P) \Delta \eta = B_{tq} \Delta \zeta + A_t^b \Delta \eta + |b_t\rangle (\Delta \theta_s + \varpi), \\ \Delta \theta_s = \langle c | \Delta \eta \end{cases} \quad (\text{A1})$$

where, for instance,

$$A_q = \frac{\partial^2 \zeta}{\partial_t \partial \zeta}, \quad B_{qt} = \frac{\partial^2 \zeta}{\partial_t \partial \eta}, \quad |b_{qt}\rangle = \frac{\partial^2 \zeta}{\partial_t \partial \theta_s}, \quad (\text{A2})$$

and similarly for all the other constant Jacobian matrices and vectors. The direct influence of the ζ variables in the η system is from $B_{tq} \Delta \zeta$. Thus, we obtain

$$(B_{tq} \Delta \zeta)_{(i)} = \frac{\partial^2 \eta_{(i)}}{\partial_t \partial \zeta_{(k)}} \Delta \zeta_{(k)}, \quad (\text{A3})$$

the summation on k running on the dimension of system ζ . With this expression, it is seen how each ζ variable at gridpoint k impacts temperatures $\Delta \eta$ at gridpoint i .

Indirect influence comes from the feedback of ζ on η in the loop $\varpi \rightarrow |b_i\rangle \rightarrow \Delta\eta \rightarrow \mathbf{B}_{qt}\Delta\zeta \rightarrow \cdots \Delta\eta$.

The algebraic elimination of ζ provides the reduced system in temperatures, the “ T system”:

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P})\Delta\eta = [\mathbf{A}_t^b + \mathbf{B}_{tq}(\mu\mathbf{I} - \mathbf{A}_q)^{-1}\mathbf{B}_{qt}]\Delta\eta + [|b_t\rangle + \mathbf{B}_{tq}(\mu\mathbf{I} - \mathbf{A}_q)^{-1}|b_{qt}\rangle](\Delta\theta_s + \varpi), \\ \Delta\theta_s = \langle c|\Delta\eta \end{cases} \quad (\text{A4})$$

which results in the new matrix system:

$$\begin{cases} (\mu\mathbf{K} + \mathbf{P})\Delta\eta = \mathbf{A}\Delta\eta + |b\rangle(\Delta\theta_s + \varpi) \\ \Delta\theta_s = \langle c|\Delta\eta \end{cases} \quad (\text{A5})$$

One thus obtains the matrices $\mathbf{A}(\mu)$ and $|b(\mu)\rangle$ in the reduced system (A5), formed with μ -rational fraction coefficients, as

$$\begin{cases} \mathbf{A} = [\mathbf{A}_t^b + \mathbf{B}_{tq}(\mu\mathbf{I} - \mathbf{A}_q)^{-1}\mathbf{B}_{qt}] \\ |b\rangle = |b_t\rangle + \mathbf{B}_{tq}(\mu\mathbf{I} - \mathbf{A}_q)^{-1}|b_{qt}\rangle \end{cases} \quad (\text{A6})$$

Note, in particular, how $|b\rangle$ shows the additive incidence of perturbation ϖ on $\Delta\eta$ coming from the other variables of the system $\Delta\zeta$. The feedback gain takes the form

$$g(\mu) = \langle c|[\mu\mathbf{K} + \mathbf{P} - \mathbf{A}_t^b + \mathbf{B}_{tq}\mathbf{M}_q(\mu)\mathbf{B}_{qt}]^{-1} \times [|b_t\rangle + \mathbf{B}_{tq}\mathbf{M}_q(\mu)|b_{qt}\rangle], \quad (\text{A7})$$

where the matrices involved now depend on the Laplace variable through $\mathbf{M}_q = (\mu\mathbf{I} - \mathbf{A}_q)^{-1}$. It is noteworthy that the original TLS partial derivatives are resulting in matrices that are no longer partial derivatives—the reason why we had to adopt a matrix formulation.

APPENDIX B

Woodbury Feedback Formula

The formal identity to put forward brings out two fundamentally different readings of the system response to perturbation. It links the response form to the gain form of the TLS characterized by its Jacobian matrix $\mathbf{M}$ that can be diversely decomposed into submatrices $\mathbf{M} = \mathbf{M}_i - |b_i\rangle\langle c|$ ($\mathbf{M}_i$ regular). This identity reads

$$q_i = \langle c|(\mathbf{M}_i - |b_i\rangle\langle c|)^{-1}|b_i\rangle = \frac{\langle c|\mathbf{M}_i^{-1}|b_i\rangle}{1 - \langle c|\mathbf{M}_i^{-1}|b_i\rangle}, \quad (\text{B1})$$

thus giving two forms for q_i when we take $\mathbf{M}_i = \mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b$. In these forms, the lhs form expresses the system response to the perturbation exciting some mechanism represented by $|b_i\rangle$. In contrast, the structure in the rhs form explicitly gives a feedback loop associated with that mechanism—with $q_i(\mu) = g_i^*(\mu)/[1 - g_i^*(\mu)]$. This

can be easily demonstrated using the Woodbury–Sherman–Morrison theorem (see the supplemental material).

If we now set $\mathbf{M}_i = (\mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b)$ and $\mathbf{M} = \mu\mathbf{K} + \mathbf{P} - \mathbf{A}^b - \sum_i |b_i\rangle\langle c|$ in the preceding development, we further establish that

$$q_i(\mu) = \frac{g_i(\mu)}{1 - g(\mu)}, \quad (\text{B2})$$

which expresses the response in terms of the exclusive gains. The following identity links the two approaches:

$$q_i(\mu) = \frac{g_i(\mu)}{1 - \sum_j g_j(\mu)} = \frac{g_i^*(\mu)}{1 - g_i^*(\mu)}. \quad (\text{B3})$$

Hence, we have two mathematically equivalent views of the system reaction: (i) from a set of p exclusive feedback loops, and (ii) from p unique feedback loops, each one corresponding to an inclusive response of the system. The consistency between the two is summarized in

$$\sum_{j=1,p} q_i = \frac{g}{1 - g}, \quad (\text{B4})$$

which shows the uniqueness of the global response ΔT_s found by both methods.

APPENDIX C

The Decomposed Feedback System

a. TOA radiative budget and sensitivity factors

Here, we establish the expressions giving the TOA radiative impacts of a change in surface temperature $T_s \rightarrow T_s + \Delta\theta_0$. We recall the notation to identify the last row and rightmost column of any matrix $\mathbf{M}$:

$$\mathbf{M} = \begin{bmatrix} \overline{\mathbf{M}}, & |\overline{m}\rangle \\ \langle \overline{m}|, & \overline{m} \end{bmatrix}. \quad (\text{C1})$$

The TOA radiative forcing response to a surface temperature perturbation of the no-feedback system; that is, setting $|b\rangle = 0$ in (A5) is, at equilibrium

$$N = \langle 1|(\mathbf{P} - \mathbf{A}^b)\Delta\eta_0, \quad (\text{C2})$$

where $\langle 1|$ is the summing operator, assuming equal-mass layers for simplicity. As $\mathbf{P}$ gives the LW response in the reference run, the LW and SW effects of the perturbed system are accounted for in $\mathbf{A}^b$, and the additional material fluxes cancel out in the summing operation, leaving only radiation TOA budgets (LW and SW). In terms of the matrix subelements, with the correspondence $|b\rangle = 0 \rightarrow |\bar{a}\rangle = 0$ and $\bar{a} = 0$, system (C2) reads

$$N = [\langle 1|[(\bar{\mathbf{P}} - \bar{\mathbf{A}}) + \langle \bar{p} - \bar{a}| \Delta \bar{\boldsymbol{\eta}}_0 + (\langle 1| \bar{p} + \bar{p}) \Delta \theta_0]. \quad (\text{C3})$$

Although the surface layer is not in equilibrium because it is perturbed, the troposphere, perturbed by an absorbed part of $\bar{p} \Delta \theta_0$ leaving the surface, is indeed in equilibrium after a transition period. This determines its temperature change from the equation

$$(\bar{\mathbf{P}} - \bar{\mathbf{A}}) \Delta \bar{\boldsymbol{\eta}}_0 + |\bar{p}\rangle \Delta \theta_0 = 0, \quad (\text{C4})$$

yielding

$$\begin{aligned} N &= [-\langle 1| \bar{p}\rangle - \langle \bar{p} - \bar{a}| (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{p}\rangle + \langle 1| \bar{p} + \bar{p}] \Delta \theta_0 \\ &= [\bar{p} - \langle \bar{p} - \bar{a}| (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{p}\rangle] \Delta \theta_0, \end{aligned} \quad (\text{C5})$$

thus giving the surface Planck feedback factor

$$-\lambda_0 = \bar{p} - \langle \bar{p} - \bar{a}| (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{p}\rangle. \quad (\text{C6})$$

In the same manner, the λ_i are found as

$$\begin{aligned} \lambda_i &= \bar{a}_i - \langle \bar{p} - \bar{a}| (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{a}_i\rangle \\ -(\lambda_0 + \lambda) &= \bar{p} - \bar{a} - \langle \bar{p} - \bar{a}| (\bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{p} - \bar{a}\rangle. \end{aligned} \quad (\text{C7})$$

b. Feedback gain in the decomposed system

Let us consider the original system (26), which focuses on the i th mechanism,^{C1}

$$\begin{cases} (\mu \mathbf{K} + \mathbf{P} - \mathbf{A}^i) \Delta \boldsymbol{\eta} = |b_i\rangle (\Delta \theta_i + \Delta \theta_0) \\ \Delta \theta_i = \langle c| \Delta \boldsymbol{\eta} \end{cases} \quad (\text{C8})$$

and rewrite it using the matrix subelements as in (C3):

$$\begin{cases} (\mu \bar{\mathbf{K}} + \bar{\mathbf{P}} - \bar{\mathbf{A}}) \Delta \bar{\boldsymbol{\eta}} + |\bar{p} - \bar{a}^i\rangle \Delta \eta_n = |\bar{a}_i\rangle (\Delta \theta_i + \Delta \theta_0) \\ \langle \bar{p} - \bar{a}| \Delta \bar{\boldsymbol{\eta}} + (\mu \bar{k} + \bar{p} - \bar{a}^i) \Delta \eta_n = \bar{a}_i (\Delta \theta_i + \Delta \theta_0) \\ \Delta \theta_i = \Delta \eta_n; \end{cases} \quad (\text{C9})$$

with the notation $\bar{a}^i = \bar{a} - \bar{a}_i$. This way of writing the feedback system implies that $\Delta \eta_n$, the n th temperature component, identifies with the effective-response $\Delta \theta_i = \langle c| \Delta \boldsymbol{\eta} = \Delta \eta_n$ with our convention. Eliminating $\Delta \bar{\boldsymbol{\eta}}$ from the second line of the first equation and $\Delta \eta_n$ from the third, we obtain the form

$$(1 - g_i^*) \Delta \theta_i = g_i^* \Delta \theta_0, \quad (\text{C10})$$

from which it is seen that the feedback gain, the same as in Eq. (29), is now expressed as

$$g_i^*(\mu) = \frac{\bar{a}_i - \langle \bar{p} - \bar{a}| (\mu \bar{\mathbf{K}} + \bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{a}_i\rangle}{\mu \bar{k} + \bar{p} - \bar{a}^i - \langle \bar{p} - \bar{a}| (\mu \bar{\mathbf{K}} + \bar{\mathbf{P}} - \bar{\mathbf{A}})^{-1} |\bar{p} - \bar{a}^i\rangle}, \quad (\text{C11})$$

which identifies with $-\lambda_i/(\lambda_T - \lambda_i)$ at equilibrium. We also obtain a relation between sensitivities and effective and forcing surface temperature perturbations:

$$-\frac{\Delta \theta_i}{\lambda_i} = \frac{\Delta \theta_0}{\lambda_T - \lambda_i}. \quad (\text{C12})$$

For reference, the dynamical counterparts of TOA sensitivities are found to be

$$\begin{aligned} -\lambda_0(\mu) &= \mu \bar{k} + \bar{p} - \langle -\mu \bar{k} + \bar{p} - \bar{a}(\mu)| [\bar{\mathbf{P}} - \bar{\mathbf{A}}(\mu)]^{-1} |\bar{p}\rangle \\ \lambda_i(\mu) &= \bar{a}_i(\mu) - \langle -\mu \bar{k} + \bar{p} - \bar{a}(\mu)| [\bar{\mathbf{P}} - \bar{\mathbf{A}}(\mu)]^{-1} |\bar{a}_i(\mu)\rangle, \end{aligned} \quad (\text{C13})$$

with the notation $\langle \bar{k}| = [k_1, k_2, \dots, k_{n-1}]$.

REFERENCES

- Arrhénius, S., 1896: On the influence of carbonic acid in the air upon the temperature of the ground. *Philos. Mag. J. Sci.*, **41**, 237–276.
- Bode, H., 1945: *Network Analysis and Feedback Amplifier Design*. Van Nostrand, 551 pp.
- Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate*, **19**, 3445–3482.
- Cai, M., and J. Lu, 2009: A new framework for isolating individual feedback processes in coupled general circulation climate models. Part II: Method demonstrations and comparisons. *Climate Dyn.*, **32**, 887–900, doi:10.1007/s00382-008-0424-4.
- Cess, R., and Coauthors, 1989: Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513–516.
- Coakley, J., Jr., 1977: Feedbacks in vertical-column energy balance models. *J. Atmos. Sci.*, **34**, 465–470.
- Colman, R., 2003: A comparison of climate feedbacks in general circulation models. *Climate Dyn.*, **20** (7–8), 865–873.
- , S. Power, and B. McAvaney, 1997: Non-linear climate feedback analysis in an atmospheric general circulation model. *Climate Dyn.*, **13**, 717–731.

^{C1} Dropping the suffix r for clarity.

- Dessler, A. E., 2010: A determination of the cloud feedback from climate variations over the past decade. *Science*, **330**, 1523–1527.
- Dufresne, J., and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmospheric–ocean models. *J. Climate*, **21**, 5135–5144.
- Forster, P., and J. Gregory, 2006: The climate sensitivity and its components diagnosed from Earth Radiation Budget data. *J. Climate*, **19**, 39–52.
- , and K. Taylor, 2006: Climate forcing and climate sensitivities diagnosed from coupled climate model integrations. *J. Climate*, **19**, 6181–6194.
- Gregory, J. M., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, **31**, L03205, doi:10.1029/2003GL018747.
- Hall, A., 2004: The role of surface albedo feedback in climate. *J. Climate*, **17**, 1550–1568.
- , and S. Manabe, 1999: The role of water vapor feedback in unperturbed climate variability and global warming. *J. Climate*, **12**, 2327–2346.
- Hallegatte, S., A. Lahellec, and J.-Y. Grandpeix, 2006: An elicitation of the dynamic nature of water feedback in climate change using a 1D model. *J. Atmos. Sci.*, **63**, 1878–1894.
- Hansen, J., D. Lacis, D. Rind, G. Russell, P. Stone, I. Fung, R. Ruedy, and J. Lerner, 1984: Climate sensitivity: Analysis of feedback mechanisms. *Climate Processes and Climate Sensitivity*, *Geophys. Monogr.*, Vol. 29, Amer. Geophys. Union, 130–162.
- , M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *J. Geophys. Res.*, **102** (D6), 6831–6864.
- Hourdin, F., and Coauthors, 2013: LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dyn.*, **40** (9–10), 2193–2222, doi:10.1007/s00382-012-1343-y.
- Ingram, W. J., C. A. Wilson, and J. F. B. Mitchell, 1989: Modeling climate change: An assessment of sea ice and surface albedo feedbacks. *J. Geophys. Res.*, **94** (D6), 8609–8622.
- Knutti, R., S. Krähenmann, D. Frame, and M. R. Allen, 2008: Comment on “Heat capacity, time constant, and sensitivity of Earth’s climate system” by S. E. Schwartz. *J. Geophys. Res.*, **113**, D15103, doi:10.1029/2007JD009473.
- Lahellec, A., S. Hallegatte, J.-Y. Grandpeix, P. Dumas, and S. Blanco, 2008: Feedback characteristics of nonlinear dynamical systems. *Euro. Phys. Lett.*, **81**, 6001–6006.
- LeTreut, H., Z. Li, and M. Forichon, 1994: Sensitivity of the LMD general circulation model to greenhouse forcing associated with two different cloud water parameterizations. *J. Climate*, **7**, 1827–1841.
- Lu, J., and M. Cai, 2009: A new framework for isolating individual feedback processes in coupled general circulation climate models. Part I: Formulation. *Climate Dyn.*, **32**, 873–885, doi:10.1007/s00382-008-0425-3.
- Manabe, S., and R. Wetherald, 1975: The effects of doubling CO₂ concentration on the climate of a general circulation model. *J. Atmos. Sci.*, **32**, 3–15.
- Mitchell, J., T. Johns, M. Eagles, W. Ingram, and R. Davis, 1999: Towards the construction of climate change scenarios. *Climate Change*, **41** (3–4), 547–581.
- Morcrette, J. J., 1991: Radiation and cloud radiative properties in the ECMRWF forecasting system. *J. Geophys. Res.*, **96** (D5), 9121–9132.
- Pedlosky, J., 1982: *Geophysical Fluid Dynamics*. Springer-Verlag, 624 pp.
- Peixoto, J. P., and A. H. Oort, Eds., 1992: *Physics of Climate*. American Institute of Physics, 520 pp.
- Schlesinger, M., 1986: Equilibrium and transient climatic warming induced by increased atmospheric CO₂. *Climate Dyn.*, **1**, 35–81.
- Schneider, E., B. Kirtman, and R. Lindzen, 1999: Upper tropospheric water vapor and climate sensitivity. *J. Atmos. Sci.*, **56**, 1649–1658.
- Schwartz, S. E., 2007: Heat capacity, time constant, and sensitivity of Earth’s climate system. *J. Geophys. Res.*, **112**, D24S05, doi:10.1029/2007JD008746.
- Slingo, A., J. Pamment, R. Allan, and P. Wilson, 2000: Water vapor feedbacks in the ECMWF reanalyses and Hadley Centre Climate Model. *J. Climate*, **13**, 3080–3098.
- Soden, B., A. Broccoli, and R. Hemler, 2004: On the use of cloud forcing to estimate cloud feedback. *J. Climate*, **17**, 3661–3665.
- , I. Held, and R. Colman, 2008: Quantifying climate feedbacks using radiative kernels. *J. Climate*, **21**, 3504–3520.
- Spencer, R. W., and W. D. Braswell, 2010: On the diagnosis of radiative feedback in the presence of unknown radiative forcing. *J. Geophys. Res.*, **115**, D16109, doi:10.1029/2009JD013371.
- Stephens, G., 2005: Cloud feedbacks in the climate: A critical review. *J. Climate*, **18**, 237–273.
- Wetherald, R., and S. Manabe, 1988: Cloud feedback processes in a general circulation model. *J. Atmos. Sci.*, **45**, 1397–1416.
- Zhang, M., J. Hack, J. Kiehl, and R. Cess, 1994: Diagnostic study of climate feedback processes in atmospheric general circulation models. *J. Geophys. Res.*, **99** (D3), 5525–5537.