Clim Dyn (2009) 32:873–885
DOI 10.1007/s00382-008-0425-3

A new framework for isolating individual feedback processes in coupled general circulation climate models. Part I: formulation

Jianhua Lu · Ming Cai

Received: 14 November 2007 / Accepted: 14 May 2008 / Published online: 12 June 2008

Abstract This paper proposes a coupled atmosphere–surface climate feedback–response analysis method (CFRAM) as a new framework for estimating climate feedbacks in coupled general circulation models with a full set of physical parameterization packages. The formulation of the CFRAM is based on the energy balance in an atmosphere–surface column. In the CFRAM, the isolation of partial temperature changes due to an external forcing or an individual feedback is achieved by solving the linearized infrared radiation transfer model subject to individual energy flux perturbations (external or due to feedbacks). The partial temperature changes are addable and their sum is equal to the (total) temperature change (in the linear sense). The decomposition of feedbacks is based on the thermodynamic and dynamical processes that directly affect individual energy flux terms. Therefore, not only those feedbacks that directly affect the TOA radiative fluxes, such as water vapor, clouds, and ice-albedo feedbacks, but also those feedbacks that do not directly affect the TOA radiation, such as evaporation, convections, and convergence of horizontal sensible and latent heat fluxes, are explicitly included in the CFRAM. In the CFRAM, the feedback gain matrices measure the strength of individual feedbacks. The feedback gain matrices can be estimated from the energy flux perturbations inferred from individual parameterization packages and dynamical modules. The inter-model spread of a feedback gain matrix would help us to detect the origins of the uncertainty of future climate projections in climate model simulations.

J. Lu · M. Cai (✉)
Department of Meteorology,
Florida State University, Tallahassee, FL 32306, USA
e-mail: cai@met.fsu.edu

Keywords Climate feedback · Global warming · Climate sensitivity

1 Introduction

The uncertainty in projections of future climate changes and in climate feedback estimates is still a challenge to the climate modeling community for a better understanding of underlying mechanisms of climate changes and climate sensitivity. Various methods have been put forward for studying climate feedback and sensitivity and for quantifying the uncertainties in climate feedbacks and sensitivities among different climate model simulations (Hansen et al. 1984; Wetherald and Manabe 1988; Cess et al. 1990; Zhang et al. 1994; Hall and Manabe 1999; Schneider et al. 1999; Held and Soden 2000; Colman 2003; Boer and Yu 2003; Soden and Held 2006; Hall and Qu 2006; Winton 2006; Bony et al. 2006).

It is understood that every climate variable that responds to global surface temperature change and affects the earth's radiation budget could constitute a climate feedback agent (Bony et al. 2006). However, the commonly used framework for climate feedback analysis, based on the premise that the temperature change is the response to (radiative) energy exchange with outer space, has focused on the radiation perturbation at top of the atmosphere (TOA), or at the tropopause for troposphere–surface system. See Ramaswamy et al. (2001). Accordingly, the climate feedback agents are those mechanisms that directly affect the radiative budget at TOA, namely, water vapor, cloud, surface albedo, and atmospheric temperature. These feedback agents are considered in the "partial radiative perturbation" (PRP) method (Wetherald and Manabe 1988), in the "cloud forcing analysis method" (CRF, Cess et al. 1990), and in the "online feedback suppression method" (Hall and Manabe 1999; Schneider et al. 1999). Readers may refer to Bony et al. (2006) for a thorough review on the strengths and limitations of these three methods (see Soden et al. (2004) for a comparison between the PRP and CRF and Stephens (2005) for a review on cloud feedback analysis).

The existing feedback analysis methods do not provide a direct estimate of the addable contributions to the total temperature change from individual feedback agents. In the PRP method, the strength of a feedback agent is measured by its feedback parameter, defined as the ratio of the total radiative perturbation at the TOA due to a specific feedback agent to the total change of the surface temperature. As a result, one cannot use the feedback parameter to estimate the partial surface temperature change due to the change caused by an individual feedback agent. The online feedback suppression method, designed to isolate the individual contribution from a single feedback agent, can provide an estimate of the partial surface temperature change due to the specific feedback agent. As to be shown in Part II of this series of papers, the partial temperature changes estimated by the online feedback suppression method include the "compensating effects" of other feedbacks when a specific feedback is suppressed. As a result, one cannot add these partial temperature changes obtained by running the same climate model multiple times with the exclusion of only one specific feedback agent at one time to compare the total temperature change obtained with all feedback agents included.

It is of importance to point out that there are no universal definitions for climate forcing, feedback, and response. These three concepts are inseparable and the definitions of the latter two depend on the climate variable under the consideration and the definition of climate forcing. In general, a feedback is defined as "induced input from the output". In the PRP method, only the surface temperature change is regarded as the output of the system in response to an external forcing. Then feedbacks are the subsequent energy flux perturbations of the climate system ("inputs") which all are assumed to be directly and indirectly induced by the change in the surface temperature (the "output"). Physically speaking, all changes in the climate system, including atmosphere and surface temperatures, water vapor, cloud, precipitation, convections, atmospheric and oceanic circulations, are system responses to an external forcing. The feedbacks are the energy flux perturbations ("inputs") induced by the system responses collectively, not necessarily only by any specific variable individually (say the surface temperature change). This is the key difference between the climate system and an electrical signal control system for which the concept of feedback was first adopted (Bode 1945; also see Bates 2007 for a complete survey and theoretic considerations on the concept of feedbacks in climate research). Although we cannot examine "the effects of induced inputs from the output on the output itself" based on the classic feedback analysis framework for the Earth climate system, we can nevertheless examine the "effects of induced inputs from all outputs (system responses) on one specific output alone". For the Earth climate system, the "induced inputs" are the energy flux perturbations associated with individual physical and dynamical processes and one of the "specific outputs" is temperature. As to be shown in the remainder of the paper, this generalized concept of the feedback and response enables us to calculate the partial temperature changes in response to the energy flux perturbations caused by the external forcing and subsequent feedbacks. These partial temperature changes measure the strength of individual feedbacks. Moreover, these partial changes are addable and can be directly compared to the (total) temperature change.

In a TOA-based climate feedback analysis (e.g., PRP), the induced "inputs" are defined as the energy flux perturbations at the TOA only, as the external forcing. In that regard, the thermodynamic and dynamical processes that do not directly influence the TOA energy balance, such as surface evaporation, surface sensible heat flux, and vertical convections, are not "feedbacks" because they cannot contribute to "inputs". However, the external forcing has a vertical and horizontal structure, and it is its 3D heating/cooling perturbation that causes the climate change. In this sense, the changes in energy exchanges between the atmosphere and surface, and in horizontal and vertical redistribution of energy due to changes in atmospheric and oceanic motions ("outputs") can be regarded as "inputs" or feedbacks because they do act to either strengthen, or weaken, or even oppose the 3D external forcing, although they may not necessarily cause a radiative perturbation at the TOA. Therefore, we ought to generalize the concept of the TOA-based climate forcing and climate feedbacks to 3D climate forcing and climate feedbacks. This leads to the other main difference between the method to be formulated in this paper and a TOA-based method.

In light of the discussions above, we here propose the following definitions of climate forcing, climate feedbacks, and climate response. *Climate forcing* is the change in the vertical difference in net radiative energy flux (Wm^{-2}) in the atmosphere and the change in net radiative energy flux (Wm^{-2}) at the surface, due to a change in the external factors of the climate system, such as a change in CO_2, aerosols, or the solar energy flux. *Climate feedback* is defined as the subsequent changes in both non-temperature induced radiative and non-radiative energy fluxes of the climate system that either strengthen, or weaken, or even oppose the original climate forcing. The non-temperature induced radiative energy flux perturbations include the changes in the vertical difference in net radiative energy flux (Wm^{-2}) in the atmosphere and the change in net radiative energy flux at the surface, due to changes in the water vapor, cloud, or surface albedo (thermodynamic feedbacks). The non-radiative energy flux perturbations include (a) the changes in the vertical and horizontal difference (Wm^{-2}) of the transport of total energy (the sum of moist static energy and kinetic energy) in the atmosphere and the changes in the surface turbulent energy exchanges with the atmosphere, and (b) the changes in horizontal difference (Wm^{-2}) of the oceanic transport of total (oceanic) energy in the surface column. *Climate response* is the changes in temperatures of atmosphere and surface in response to the climate forcing and climate feedbacks. For the sake of brevity, unless specified otherwise, we hereafter will refer to non-temperature induced radiative and non-radiative energy flux perturbations simply as "energy flux perturbations".

Based on the generalized definitions of climate forcing, climate feedback, and climate response, we propose in this two-part series of papers a new climate feedback analysis framework. We refer to the new framework as the coupled atmosphere–surface "climate feedback–response analysis method" (abbreviated as "CFRAM"). In Part I of this two-part series of papers, we present the mathematical formulation of the CFRAM whereas in Part II, we demonstrate how to use the CFRAM to diagnose climate feedbacks and compare the CFRAM with the PRP and online feedback suppression methods in the context of a coupled atmosphere–surface single column climate model.

A prototype approach of the new climate feedback analysis framework has already been applied in the context of a 4-box radiative–transportive climate model (Cai 2006; Cai and Lu 2007) for explaining the role of poleward heat transport in contributing to the polar warming amplification by isolating it from other local thermodynamic and dynamical processes (evaporation, water vapor, and ice-albedo feedbacks). The CFRAM is generalized from our prototype approach in aiming to evaluate the partial contributions to total temperature change from each energy transfer and transport process in the context of coupled general circulation models with a full physical parameterization package. The formulation of the CFRAM is based on the energy balance in both the atmosphere and the land/ocean column underneath. We take advantage of the fact that the infrared radiation is explicitly and directly related to temperatures in the entire atmosphere–surface column. Therefore, the temperature changes in the equilibrium response to any perturbation in other energy flux terms, external or due to feedbacks, can be determined by requiring the temperature-induced change in the infrared radiation to exactly balance the non-temperature induced energy flux perturbations.

In the CFRAM, the isolation of coupled atmosphere–surface responses to individual feedbacks is achieved by solving the linearized infrared radiation transfer model subject to climate forcing and individual climate feedbacks. The decomposition of feedbacks is based on the thermodynamic and dynamical processes that directly represent individual energy flux terms in the energy balance equations for the atmosphere and the surface. Because the changes in air and surface temperatures are calculated simultaneously, the (air) temperature or lapse rate feedback, by definition, no longer exists in the CFRAM. Furthermore, the isolated responses to the external forcing alone or individual feedbacks are additive and their sum is the total response to the external forcing.

It is of importance to add here that the isolation of the contributions to the atmosphere and surface temperature change from climate forcing and climate feedbacks using the CFRAM is not based on the assumption that the individual processes are physically independent with each other. It is fully understood that the climate forcing and all individual feedback processes work synergistically under the thermodynamic and dynamic constraints intrinsic to the energy cycle within the climate system. In the CFRAM, we do not concern how the interactive relations among individual feedback processes and temperature changes give rise to the energy flux perturbations although in the PRP, it is explicitly assumed that the energy flux perturbations of individual feedbacks are caused by the change in the surface temperature. For given perturbation climate simulations, we should be able to know all changes in energy cycle due to individual thermodynamic and dynamical processes (provided that we output them from the climate perturbation simulation). We then can apply the CFRAM, as an offline diagnostic tool, to evaluate the partial temperature change associated with an individual energy flux perturbation by requiring the infrared radiation induced by the temperature change alone to exactly balance the energy flux perturbation under consideration. The sum of these partial temperature changes independently calculated by the CFRAM would correspond to the (total) temperature of the new equilibrium climate state in response to the climate forcing. In this sense, we "isolate" the contributions to the (total) temperature change from the external forcing alone, and from individual feedbacks although physically speaking, these feedbacks are not independent with one another.

The organization of the presentation is as follows. Presented in Sect. 2 is the mathematical formulation of the CFRAM. In Sect. 3, we derive the TOA version of the CFRAM (abbreviated as TFRAM hereafter) to illustrate the differences in the mathematical formulations of the CFRAM and PRP methods. Section 4 discusses the feedback gains in the TFRAM and feedback gain matrices in the CFRAM to illustrate some subtle but intrinsic differences between the PRP and CFRAM methods. In Sect. 5, we show a derivation confirming that the lapse rate feedback defined in a TOA-based approach consists of individual contributions from the external forcing, and from each of physical and dynamical processes in the climate system. A brief summary about the main features of the CFRAM is provided in Sect. 6.

2 The mathematical formulation of the CFRAM

The mathematical formulation of the CFRAM is based on the conservation equation of total energy (Peixoto and Oort 1992). Total energy $\mathscr{E}$ in the atmosphere is defined as the sum of dry static energy, latent heat, and kinetic energy, $\mathscr{E} = c_pT + gz + Lq + \frac{1}{2}\vec{V}\cdot\vec{V}$, where T, q, $\vec{V}$ are atmospheric temperature, specific humidity, and 3D velocity fields, respectively; z is the height; c_p, g, and L are air specific heat constant at constant pressure, gravity, and latent heat constant, respectively. Let the atmosphere be divided into M layers with the convention that the first layer represents the top layer of the atmosphere and the surface column (either land or ocean) as the $(M + 1)$th layer. We only consider the equilibrium state by applying a long time average to all terms in the equation of total energy. In the presentation below, we omit the symbols for the time mean and horizontal location indices for the simplicity. In each layer, all terms in the energy equation have a unit of Wm^{-2}. At a given horizontal location, the time mean energy balance equation in the atmosphere–surface column is

$$\vec{\mathbf{R}} = \vec{\mathbf{S}} + \vec{\mathbf{Q}}^{conv} + \vec{\mathbf{Q}}^{turb} - \vec{\mathbf{D}}^{v} - \vec{\mathbf{D}}^{h} + \vec{\mathbf{W}}^{fric}, \text{or},$$

$$\begin{pmatrix} R_1 \\ \vdots \\ R_M \\ R_{M+1} \end{pmatrix} = \begin{pmatrix} S_1 \\ \vdots \\ S_M \\ S_{M+1} \end{pmatrix} + \begin{pmatrix} Q_1^{turb} \\ \vdots \\ Q_M^{turb} + LE + H \\ -LE - H \end{pmatrix} + \begin{pmatrix} Q_1^{conv} \\ \vdots \\ Q_M^{conv} \\ 0 \end{pmatrix} - \begin{pmatrix} D_1^{v} \\ \vdots \\ D_M^{v} \\ 0 \end{pmatrix} - \begin{pmatrix} D_1^{h} \\ \vdots \\ D_M^{h} \\ D_{M+1}^{h} \end{pmatrix} + \begin{pmatrix} W_1^{fric} \\ \vdots \\ W_M^{fric} \\ W_{M+1}^{fric} \end{pmatrix} \quad (1)$$

where $\mathbf{R} = (R_1, \ldots, R_M, R_{M+1})^T$ is the energy flux vector whose elements are the net infrared radiation flux leaving the mth layer atmosphere for m = 1, 2, ... M and the net infrared radiation leaving the surface layer for $m = M + 1$; S_m in $\vec{\mathbf{S}} = (S_1, \ldots, S_M, S_{M+1})^T$ is the solar radiation flux absorbed by the mth layer atmosphere for $m \le M$ and S_{M+1} solar radiation flux absorbed at land (or ocean) surface (note that $S_{M+1} = S^{\downarrow}_{M+1}(1 - \alpha_s)$, where $S^{\downarrow}_{M+1}$ the downward solar radiation flux at the surface and α_s is the surface albedo); Q_m^{turb} in $\vec{\mathbf{Q}}^{turb} = (Q_1^{turb}, \ldots, Q_M^{turb} + LE + H, -LE - H)^T$ is the convergence of total energy $\mathscr{E}$ into the mth layer due to the turbulent motions (also including the numerical diffusion); The turbulent sensible (H) and latent heat (LE) transfer from the surface to the atmosphere is included as part of the turbulent heating (cooling) to the bottom of the atmosphere (the surface column); Q_m^{conv} in $\vec{\mathbf{Q}}^{conv} = (Q_1^{conv}, \ldots, Q_M^{conv}, 0)^T$ is the convergence of total energy into the mth layer from other layers due to convective motions; D_m^v in $\vec{\mathbf{D}}^{v} = (D_1^v, \ldots, D_M^v, 0)^T$ is the large-scale vertical transport of total energy out of the mth layer into other layers at the same horizontal location; D_m^h in $\vec{\mathbf{D}}^{h} = (D_1^h, \ldots, D_M^h, D_{M+1}^h)^T$ for $m \le M$ is the horizontal transport of total energy out of the horizontal box into its neighbor boxes at the same mth layer and $D_{M+1}^h = 0$ if the surface column is land and otherwise $D_{M+1}^h = D_o$ is the horizontal energy flux out of the ocean column into its neighbor ocean columns; the work done by the atmospheric friction force is given by the first M elements in $\vec{\mathbf{W}}^{fric} = (W_1^{fric}, \ldots, W_M^{fric}, W_{M+1}^{fric})^T$ whereas the last element $W_{M+1}^{fric} = 0$ if the surface column is land and otherwise $W_{M+1}^{fric} = W_0^{fric}$ is the net energy input due to the work done by the surface wind stress. Readers may consult with Peixoto and Oort (1992) for details of these energy flux terms. It should be noted that the vertical summation of all elements in each of $\vec{\mathbf{Q}}^{conv}$, $\vec{\mathbf{Q}}^{turb}$, $\vec{\mathbf{D}}^{v}$, and $\vec{\mathbf{W}}^{fric}$ is equal to zero. The vertical summation of $\vec{\mathbf{D}}^{h}$ is not equal to zero and only the global and vertical mean of $\vec{\mathbf{D}}^{h}$ is equal to zero.

Now let us consider the climate response to an external perturbation forcing

$$\Delta\vec{\mathbf{F}}^{ext} = (F_1^{ext}, \ldots, F_M^{ext}, F_{M+1}^{ext})^T \quad (2)$$

which may be a radiation forcing induced by anthropogenic greenhouse gases, or by aerosols, or by a change in Ozone, or by a change in incident solar radiation. In response to the external perturbation forcing, the climate state changes to a new equilibrium state. The difference in the energy flux terms between the new and unperturbed equilibrium states satisfies

$$\Delta\vec{\mathbf{R}} = \Delta\vec{\mathbf{F}}^{ext} + \Delta\vec{\mathbf{S}} + \Delta\vec{\mathbf{Q}}^{conv} + \Delta\vec{\mathbf{Q}}^{turb} - \Delta\vec{\mathbf{D}}^{v} - \Delta\vec{\mathbf{D}}^{h} + \Delta\vec{\mathbf{W}}^{fric} \quad (3)$$

where "Δ" stands for the difference between the two equilibrium states. Using the linear approximation, the change in the infrared radiative flux vector ($\Delta\vec{\mathbf{R}}$) can be further decomposed into

$$\Delta\vec{\mathbf{R}} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)\Delta\vec{\mathbf{T}} + \Delta^{(w)}\vec{\mathbf{R}} + \Delta^{(c)}\vec{\mathbf{R}} \tag{4}$$

where $\Delta^{(c)}\vec{\mathbf{R}}$ and $\Delta^{(w)}\vec{\mathbf{R}}$ are the changes in the infrared radiative fluxes due to changes in water vapor and clouds, respectively. The term $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)\Delta\vec{\mathbf{T}}$ represents the changes in the infrared radiative fluxes due to $\Delta\vec{\mathbf{T}} = (\Delta T_1, \ldots, \Delta T_M, \Delta T_{M+1})^T$, temperature changes throughout the entire atmosphere–surface column. The matrix $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)$ is

$$\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right) = \begin{pmatrix} \frac{\partial R_1}{\partial T_1} & \cdots & \frac{\partial R_1}{\partial T_{M+1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial R_{M+1}}{\partial T_1} & \cdots & \frac{\partial R_{M+1}}{\partial T_{M+1}} \end{pmatrix} \tag{5}$$

The matrix $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)$ is equivalent to the Planck feedback parameter in the PRP method or the Stefan–Boltzmann feedback factor in a zero-dimensional climate model. We will refer to $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)$ as the Planck feedback matrix hereafter.

The changes in solar radiation fluxes ($\Delta\vec{\mathbf{S}}$) can be expressed as the sum of that due to surface-albedo feedback ($\Delta^{(\alpha)}\vec{\mathbf{S}}$), due to cloud-albedo feedback ($\Delta^{(c)}\vec{\mathbf{S}}$), and due to water vapor feedback ($\Delta^{(w)}\vec{\mathbf{S}}$), or,

$$\Delta\vec{\mathbf{S}} = \Delta^{(c)}\vec{\mathbf{S}} + \Delta^{(\alpha)}\vec{\mathbf{S}} + \Delta^{(w)}\vec{\mathbf{S}} \tag{6}$$

Substituting (4) and (6) into (3), we obtain

$$\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)\Delta\vec{\mathbf{T}} = \Delta\vec{\mathbf{F}}^{ext} + \Delta^{(\alpha)}\vec{\mathbf{S}} + \Delta^{(c)}(\vec{\mathbf{S}} - \vec{\mathbf{R}}) + \Delta^{(w)}(\vec{\mathbf{S}} - \vec{\mathbf{R}}) + \Delta\vec{\mathbf{Q}}^{conv} + \Delta\vec{\mathbf{Q}}^{turb} - \Delta\vec{\mathbf{D}}^{v} - \Delta\vec{\mathbf{D}}^{h} + \Delta\vec{\mathbf{W}}^{fric} \tag{7}$$

Each term on the right hand side (RHS) of (7) represents an energy flux perturbation. Other than the external forcing, $\Delta\vec{\mathbf{F}}^{ext}$, the remaining terms are energy flux perturbations due to various feedback processes. The left hand side (LHS) of (7) is the infrared radiative energy flux perturbation due to the coupled atmosphere–surface temperature response to the energy flux perturbation terms, which balances exactly (in the linear sense) the energy flux perturbations on the RHS.

It is of importance to note that (4) and (6) are obtained with an implicit assumption that radiative perturbations can be linearized by omitting the higher order terms of each thermodynamic feedback and the interactions among the thermodynamic feedbacks (water vapor, ice-feedback, cloud feedback), as commonly adopted in the PRP method (Bony et al. 2006). The energy flux perturbation terms due to changes in atmospheric (convective, turbulent, and large-scale) motions, $\Delta\vec{\mathbf{Q}}^{conv}$, $\Delta\vec{\mathbf{Q}}^{turb}$, $\Delta\vec{\mathbf{D}}^{v}$, $\Delta\vec{\mathbf{D}}^{h}$, and $\Delta\vec{\mathbf{W}}^{fric}$ in (7) are not subject to the linearization approximation. Although these dynamic processes are ultimately coupled with one another and with the thermodynamic processes, their contributions to the energy flux perturbations can be estimated from the model outputs independently.

It follows that one can solve for temperature perturbations according to

$$\begin{aligned}\Delta\vec{\mathbf{T}} &= \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}\Big\{\Delta\vec{\mathbf{F}}^{ext} + \Delta^{(\alpha)}\vec{\mathbf{S}} + \Delta^{(c)}(\vec{\mathbf{S}} - \vec{\mathbf{R}}) + \Delta^{(w)} \\ &\qquad (\vec{\mathbf{S}} - \vec{\mathbf{R}}) + \Delta\vec{\mathbf{Q}}^{conv} + \Delta\vec{\mathbf{Q}}^{turb} - \Delta\vec{\mathbf{D}}^{v} - \Delta\vec{\mathbf{D}}^{h} + \Delta\vec{\mathbf{W}}^{fric}\Big\} \\ &= \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}\sum_{n=0}^{8}\Delta\vec{\mathbf{F}}^{(n)}\end{aligned} \tag{8}$$

where $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}$ is the inverse of the matrix $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)$; $\Delta\vec{\mathbf{F}}^{(0)}$ is the external forcing $\Delta\vec{\mathbf{F}}^{ext}$; $\Delta\vec{\mathbf{F}}^{(n)}$ $(n > 0)$ is denoted as one of the energy flux perturbation terms due to feedbacks on the RHS of (7). Obviously, the solution of (8), namely, $\Delta\vec{\mathbf{T}}$, is the vertical profile of temperature perturbations in the entire atmosphere-column whose infrared radiation perturbation balances exactly with the energy flux perturbations at each layer, at the TOA, and at the surface. For this reason, (8) is named as the coupled atmosphere–surface climate feedback–response analysis method (CFRAM).

The transition from (7) to (8) is the turning point that separates our method from the classic PRP method, besides using the TOA versus vertically varying external radiative forcing. If one would adopt the feedback concept adopted in the PRP method, one would assume that $\Delta\vec{\mathbf{F}}^{(n)}$ $(n > 0)$ on the RHS of (7) are "caused" by the change in atmosphere–surface temperature $\Delta\vec{\mathbf{T}}$. Assume that the relation between changes in $\Delta\vec{\mathbf{F}}^{(n)}$ $(n > 0)$ with the changes of temperature could be identified, one could write $\Delta\vec{\mathbf{F}}^{(n)} = \left(\frac{\partial\vec{\mathbf{F}}^{(n)}}{\partial\vec{\mathbf{T}}}\right)\Delta\vec{\mathbf{T}}$ for $n > 0$. One would then move these terms to the LHS of (7) and solver for $\Delta\vec{\mathbf{T}}$ as a total atmosphere–surface column temperature response to the external forcing $\Delta\vec{\mathbf{F}}^{ext}$.

Our method, namely solving for $\Delta\vec{\mathbf{T}}$ using (8), effectively treat energy flux perturbations due to feedback processes parallel to the energy flux perturbation due to the external forcing. It is understood that all energy flux perturbation terms due to feedbacks and energy flux perturbation due to temperature changes are coupled together and are ultimately caused by the presence of the external forcing. Particularly, some of feedbacks, such as water vapor and ice-albedo feedbacks, are explicitly and directly related to temperature changes. However, from the perspective of energy balance, these energy flux perturbations due to changes in other climate variables are indeed the energy sources/sinks for the additional temperature changes on top of the change due to the external forcing directly. In these sense, what we obtain from (8) indeed are the total temperature changes (in the linear framework) in response to the external forcing.

Because (8) is a linear equation, we can apply the linear decomposition principle to obtain the temperature response to an individual energy flux perturbation term by solving

$$\Delta\vec{\mathbf{T}}^{(n)} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} \Delta\vec{\mathbf{F}}^{(n)} \tag{9}$$

where $\Delta\vec{\mathbf{F}}^{(n)}$ is one of the energy flux perturbation terms on the RHS of (8) and $\Delta\vec{\mathbf{T}}^{(n)}$ is the coupled atmosphere–surface temperature response to that specific energy flux perturbation. Particularly, $\Delta\vec{\mathbf{T}}^{(0)} \left(\Delta\vec{\mathbf{T}}^{(0)} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} \Delta\vec{\mathbf{F}}^{ext}\right)$ is the coupled atmosphere-surface temperature response to the external forcing $\Delta\vec{\mathbf{F}}^{ext}$ alone in the absence of any feedbacks. For an energy flux perturbation term associated with a specific feedback process $\Delta\vec{\mathbf{F}}^{(n)} (n > 0), \Delta\vec{\mathbf{T}}^{(n)}$ is the coupled atmosphere–surface temperature change in response to the energy flux perturbation due to the feedback process $\Delta\vec{\mathbf{F}}^{(n)}$. In terms of the generalized concept of feedback and response, $\Delta\vec{\mathbf{T}}^{(n)}$ is the effect of an induced energy flux perturbation $\Delta^{(n)}$ from the system responses on the temperature alone. By the linear decomposition principle, the sum of all responses, namely,

$$\Delta\vec{\mathbf{T}}^{tot} = \sum_n \Delta\vec{\mathbf{T}}^{(n)} \tag{10}$$

is the total response to the external forcing after taking all feedback processes into consideration.

These feedbacks can be loosely grouped into three groups:

(1) *Radiation-related thermodynamic feedbacks* include the solar radiation perturbation due to changes in surface albedo ($\Delta^{(\alpha)}\vec{\mathbf{S}}$), the radiation perturbations due to changes in clouds ($\Delta^{(c)}(\vec{\mathbf{S}} - \vec{\mathbf{R}})$), and the perturbations due to changes in water vapor ($\Delta^{(w)}(\vec{\mathbf{S}} - \vec{\mathbf{R}})$). The coupled atmosphere–surface temperature response to the radiation-related thermodynamic feedbacks, denoted with the superscript "rad", is

$$\Delta\vec{\mathbf{T}}^{rad} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} \left(\Delta^{(\alpha)}\vec{\mathbf{S}} + \Delta^{(c)}(\vec{\mathbf{S}} - \vec{\mathbf{R}}) + \Delta^{(w)}(\vec{\mathbf{S}} - \vec{\mathbf{R}})\right) \tag{11}$$

(2) *Local dynamical feedbacks* include feedbacks due to changes in convective energy transport ($\Delta\vec{\mathbf{Q}}^{conv}$), due to the changes in turbulent (and diffusive) energy transfer in the atmosphere and at the surface ($\Delta\vec{\mathbf{Q}}^{turb}$), due to the changes in large-scale vertical energy transport ($-\Delta\vec{\mathbf{D}}^{v}$), and due to the changes in the work done by the friction ($\Delta\vec{\mathbf{W}}^{fric}$). The temperature response to the local dynamical feedbacks, denoted with the superscript "loc-dyn", is

$$\Delta\vec{\mathbf{T}}^{loc-dyn} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} \left(\Delta\vec{\mathbf{Q}}^{conv} + \Delta\vec{\mathbf{Q}}^{turb} - \Delta\vec{\mathbf{D}}^{v} + \Delta\vec{\mathbf{W}}^{fric}\right) \tag{12}$$

Note that the response to the evaporation feedback is part of $\Delta\vec{\mathbf{T}}^{loc-dyn}$ in response to $\Delta\vec{\mathbf{Q}}^{turb}$.

(3) *Non-local dynamical feedbacks* are due to the changes in horizontal energy transport by both atmospheric and oceanic motions ($-\Delta\vec{\mathbf{D}}^{h}$), responsible for exporting climate sensitivity from one region (say low latitudes) to other regions (say high latitudes) as shown in Cai (2006) and Cai and Lu (2007). The temperature response to the non-local dynamical feedbacks, denoted with the superscript "dyn", is

$$\Delta\vec{\mathbf{T}}^{dyn} = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} (-\Delta\vec{\mathbf{D}}^{h}) \tag{13}$$

Then the total response to the external forcing is

$$\Delta\vec{\mathbf{T}}^{tot} = \Delta\vec{\mathbf{T}}^{(0)} + \Delta\vec{\mathbf{T}}^{rad} + \Delta\vec{\mathbf{T}}^{loc-dyn} + \Delta\vec{\mathbf{T}}^{dyn} \tag{14}$$

Obviously, each group can be further divided into individual feedback agents.

Now let us discuss the invertibility of the Planck feedback matrix. Because the temperature responses to the external forcing and to each of the feedback terms is subject to its inverse matrix $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}$, the feasibility of the feedback–response calculation based on the CFRAM requires that the Planck feedback matrix must not be degenerated, or $\left|\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right|$ must not be close to zero, which is the necessary condition to invert $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)$ to $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}$. The terms on the diagonal of the matrix $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \vec{\mathbf{T}}}\right)$ are $\{\frac{\partial R_i}{\partial T_i}, i = 1, 2, \ldots, M, M+1\}$, representing the changes of infrared radiation flux divergence in each layer with respect to the layer temperature. The elements in the lower triangle part of the matrix are $\frac{\partial R_i}{\partial T_j}$ with $i > j$, representing the influence of changes in upper layer temperatures to the infrared radiation flux divergence at lower layers. Similarly, the elements in the upper triangle part of the matrix, $\frac{\partial R_i}{\partial T_j}$ with $i < j$, represent the influence of changes in lower layer temperatures to the infrared radiation flux divergence at upper layers. It follows that by the nature of infrared radiation transfer, the diagonal part of the matrix $\frac{\partial R_i}{\partial T_i}$ must be positive. Also an element in the diagonal part is larger than the terms in the same row and column, namely, $\left|\frac{\partial R_i}{\partial T_i}\right| > \left|\frac{\partial R_i}{\partial T_j}\right|$ and $\left|\frac{\partial R_i}{\partial T_i}\right| > \left|\frac{\partial R_j}{\partial T_i}\right|$ for $i \neq j$. Furthermore, $\left|\frac{\partial R_i}{\partial T_j}\right| (i \neq j)$ decays rapidly as the difference between i and j increases. These general properties of the Planck feedback matrix are vividly illustrated in Fig. 1. These properties would ensure

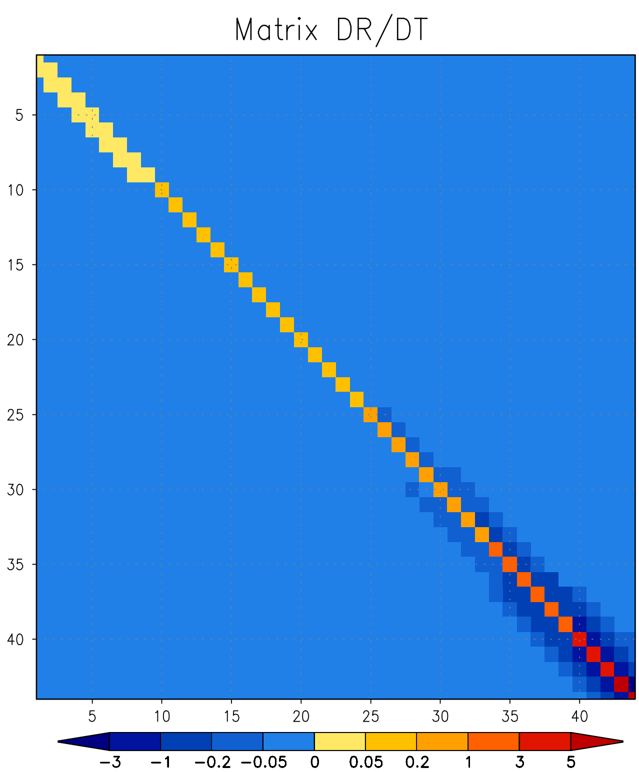

Fig. 1 Numerical values ($Wm^{-2}K^{-1}$) of the elements in a 44 × 44 Planck feedback matrix $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \vec{\mathbf{T}}}\right)$. The abscissa is the column index (j) and the ordinate the row index (i) of the matrix. The jth column of the Planck feedback matrix is the cooling rate change from the top layer ($i = 1$) to the surface layer ($i = 44$) due to a 1 K temperature increase at the jth layer from an equilibrium temperature profile of a radiative–convective model. The radiation model used in the radiative–convective model is from Fu and Liou (1993)

that the matrix $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \vec{\mathbf{T}}}\right)$ is non-singular and can be inverted to $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \vec{\mathbf{T}}}\right)^{-1}$ with high accuracy. Also since the size of the Planck matrix is the same as the number of the vertical layers in an AGCM model, which is typically in the order of less than 100, there is no computational cost issue to speak of.

3 TOA version of the CFRAM

To facilitate a direct comparison between the CFRAM and PRP methods, we need to derive the TOA version of the CFRAM (TFRAM). The TFRAM can be obtained by first reducing the coupled atmosphere–surface energy balance to a single level at the TOA after summing (7) vertically from the surface to the TOA. The vertical summation of $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \vec{\mathbf{T}}}\right)\Delta\vec{\mathbf{T}}$ is

$$\sum_{i=1}^{M+1}\left\{\sum_{j=1}^{M+1}\frac{\partial R_i}{\partial T_j}\Delta T_j\right\} = \sum_{j=1}^{M+1}\left\{\sum_{i=1}^{M+1}\frac{\partial R_i}{\partial T_j}\right\}\Delta T_j = \sum_{j=1}^{M+1}\frac{\partial R^{toa}}{\partial T_j}\Delta T_j \tag{15}$$

It follows that (7) becomes, after the vertical summation and some rearrangements,

$$\left(-\sum_{j=1}^{M+1}\frac{\partial R^{toa}}{\partial T_j}\right)\Delta T_s + \sum_{j=1}^{M}\left(-\frac{\partial R^{toa}}{\partial T_j}\right)(\Delta T_j - \Delta T_s) + \Delta^{(\alpha)}S^{toa} + \Delta^{(c)}(S^{toa} - R^{toa}) + \Delta^{(w)}(S^{toa} - R^{toa}) - \Delta D = -\Delta F^{toa} \tag{16}$$

where ΔF^{toa} is the external forcing (positive for downward) at the TOA and ΔT_s is an uniform change of the atmosphere–surface column temperature. The terms on the LHS of (16), from the left to the right, are the TOA radiation energy flux changes due to the change of the system temperature (it is called as Planck feedback, defined as the vertical homogeneous response to the external forcing, Soden and Held 2006), the lapse rate feedback, the surface albedo feedback, the cloud feedback, the water vapor feedback, and the feedback due to changes in the vertically integrated horizontal redistribution of energy by atmospheric and oceanic circulations (Cai and Lu 2007), respectively. The local dynamical feedback terms in (7), including the evaporation feedback, vanish when they are summed up vertically because the vertical redistribution of the energy in the atmosphere–land (ocean) column does not alter the energy budget at the TOA. When applying a global average, the changes in spatial redistribution of energy by atmospheric and oceanic circulations also vanish.

There are two different ways in analyzing climate feedback based on (16). The first is the PRP method.

Dividing both sides of (16) by ΔT_s (which implicitly assumes that all TOA radiative flux changes due to feedbacks are caused by the surface temperature change), we recover the feedback parameter equation in the PRP method,

$$\lambda_{tot} = \frac{-\Delta F^{toa}}{\Delta T_s} = \lambda_P + \lambda_\Gamma + \lambda_\alpha + \lambda_c + \lambda_w + \lambda_D \quad (17)$$

where,

$$\lambda_P = \left(-\sum_{j=1}^{M+1} \frac{\partial R^{toa}}{\partial T_j} \right), \quad \lambda_\Gamma = \sum_{j=1}^{M+1} \left(-\frac{\partial R^{toa}}{\partial T_j} \right) \frac{\Delta T_j - \Delta T_s}{\Delta T_s},$$
$$\lambda_\alpha = \frac{\Delta^{(\alpha)} S^{toa}}{\Delta T_s}$$
$$\lambda_c = \frac{\Delta^{(c)}(S^{toa} - R^{toa})}{\Delta T_s}, \quad \lambda_w = \frac{\Delta^{(w)}(S^{toa} - R^{toa})}{\Delta T_s},$$
$$\text{and} \quad \lambda_D = \frac{-\Delta D}{\Delta T_s} \quad (18)$$

In the literature, λ_{tot} is the total surface temperature feedback parameter in response to the external forcing; λ_P is the Planck feedback parameter and λ_Γ the lapse rate feedback parameter (the sum of Planck and lapse rate feedback parameters is referred to as the temperature feedback parameter, Bony et al. (2006)); λ_α is the surface albedo feedback parameter; λ_c is the cloud feedback parameter; λ_w is the water vapor feedback parameter; and λ_D is the (non-local) dynamical feedback parameter. As indicated in (17)–(18), the individual feedback parameters derived from the PRP method are additive. But, their effects on the surface temperature change are not addable.

The second way is to divide each term of (16) with $(-\lambda_P)$ and then to move all the terms on the LHS except the first term to RHS. This leads to the TFRAM, the TOA version of the CFRAM,

$$\Delta T_s^{tot} = \Delta T_s^P + \Delta T_s^\Gamma + \Delta T_s^\alpha + \Delta T_s^c + \Delta T_s^w + \Delta T_s^D \quad (19)$$

where,

$$\Delta T_S^P = \frac{\Delta F^{toa}}{(-\lambda_P)}, \quad \Delta T_S^\Gamma = \frac{\sum_{j=1}^{M} \left(-\frac{\partial R^{toa}}{\partial T_j} \right) (\Delta T_j - \Delta T_s)}{(-\lambda_P)},$$
$$\Delta T_S^\alpha = \frac{\Delta^{(\alpha)} S^{toa}}{(-\lambda_P)}$$
$$\Delta T_S^c = \frac{\Delta^{(c)}(S^{toa} - R^{toa})}{(-\lambda_P)}, \quad \Delta T_s^w = \frac{\Delta^{(w)}(S^{toa} - R^{toa})}{(-\lambda_P)},$$
$$\text{and} \quad \Delta T_s^D = \frac{-\Delta D}{(-\lambda_P)} \quad (20)$$

In (20), λ_P is given in the first equation of (18). Obviously, unlike in the PRP method, we have made no assumption about the dependency of the individual feedback energy flux perturbations on the (total) surface temperature change in deriving (19)–(20) from (16). We merely take these individual energy flux perturbations from the "outputs" or system responses and use them as "inputs" or feedbacks to infer their "effects" on the surface temperature changes. In parallel to the PRP method, ΔT_s^{tot} is the total surface temperature change in response to the external forcing; ΔT_s^P and ΔT_s^Γ are the surface temperature changes due to the Planck and lapse rate feedbacks, respectively; ΔT_s^α is the surface temperature change due to the surface albedo feedback; ΔT_s^c is the surface temperature change due to the cloud feedback; ΔT_s^w is the surface temperature change due to the water vapor feedback parameter; and ΔT_s^D is the surface temperature change due to the (non-local) dynamical feedback.

The number of the feedback processes considered in the TFRAM is the same as the PRP method. Furthermore, the external forcing and feedbacks in (20), namely, the terms $\Delta F^{toa}, \Delta^{(\alpha)} S^{toa}, \Delta^{(c)}(S^{toa} - R^{toa}), \Delta^{(w)}(S^{toa} - R^{toa}), \Delta D$, and the air temperature changes, ΔT_j ,$j = 1,2,...,M$, are identical to their counterparts in (18) and need to be estimated from the climate model simulation outputs. However, one of the main differences from the PRP is that the TFRAM (or CFRAM for the coupled atmosphere–surface temperature responses) enables us to explicitly calculate the temperature changes due to each of individual feedbacks from the original climate model simulations. In the PRP method, however, the "partial" temperature change in response to a specific feedback is not defined.

4 Feedback gains in the TFRAM and feedback gain matrices in the CFRAM

Some intrinsic differences between the TFRAM/CFRAM and PRP methods can also be illustrated by comparing their feedback gains. Again, the proper comparison on feedback gains can be made only between the PRP and TFRAM because feedback gains in the TFRAM become feedback gain matrices in the CFRAM.

4.1 Feedback gains in the TFRAM and comparison with the PRP feedback gains

By definition, the total surface temperature change (ΔT_s^{tot}) in response to the external forcing (ΔF^{toa}) is related to the sensitivity factor of the climate system (G) or the total feedback parameter (λ_{tot}),

$$\Delta T_s^{tot} = G\Delta F^{TOA} = (-\Delta F^{TOA}/\lambda_{tot}) \quad (21)$$

The sensitivity factor of the climate system measures the surface temperature change in response to one unit perturbation forcing at the TOA. Obviously, the values of

Table 1 Comparison of the PRP and TOA version of the CFRAM methods

	PRP	TOA CFRAM
Temperature response	$\Delta T_s^{tot} = \frac{-\Delta F^{toa}}{\lambda_{tot}} = G\Delta F^{toa}$	$\Delta T_s^{tot} = \frac{-\Delta F^{toa}}{\lambda_{tot}} = G\Delta F^{toa}$
Sensitivity factor (G)	$G = G_0 f = \frac{G_0}{1-\sum g_i} = -\frac{1}{\lambda_{tot}};\quad G_0 = -\frac{1}{\lambda_P};\quad \lambda_P = \sum_{j=1}^{M+1} \frac{\partial(S-R)}{\partial T_j} < 0$	$G = G_0 f = G_0(1 + \sum \tilde{g}_i) = -\frac{1}{\lambda_{tot}};\quad G_0 = -\frac{1}{\tilde{\lambda}_P} = -\frac{1}{\lambda_P}$
Feedback gains	$g_i = \frac{\partial(S-R)}{\partial x_i}\frac{\partial x_i}{\partial T_s}/(-\lambda_p), (i \neq P)$	$\tilde{g}_i = \frac{\partial(S-R)}{\partial x_i}\frac{\partial x_i}{\partial F^{toa}}, (i \neq P)$
Feedback parameters	$\lambda_i = (-\lambda_P) g_i, (i \neq P); \lambda_{tot} = \lambda_P + \sum_{i \neq P} \lambda_i$	$\tilde{\lambda}_i = (-\tilde{\lambda}_p)/\tilde{g}_i, (i \neq P); \frac{1}{\lambda_{tot}} = \frac{1}{\tilde{\lambda}_P} + \sum_{i \neq P} \frac{1}{\tilde{\lambda}_i}$
Temperature responses to individual feedbacks	The total temperature response is not the linear summation of that due to individual feedbacks.	$\Delta T_s^P = G_0\, \Delta F^{toa}$ and $\Delta T_s^{(i)} = G_0 \tilde{g}_i \Delta F^{toa}, (i \neq P)$; $\Delta T_s^{tot} = \Delta T_s^P + \sum_{i \neq P} \Delta T_s^{(i)}$

The subscript/superscript P represents the Planck feedback, corresponding to a vertically constant temperature change in response to the external forcing ΔF^{toa}; The subscript/superscript i represents a feedback other than Planck feedback and x_i is the corresponding feedback agent; G is the sensitivity factor, the ratio of output signal (ΔT_s^{tot}) to the input signal ΔF^{toa}; G_0 equals to the inverse of the Planck feedback factor, representing the initial sensitivity factor without any feedbacks; S is the incoming solar radiation at TOA; R is the outgoing infrared radiation at TOA; T_j is the equilibrium temperature in atmosphere ($j \leq M$) and at the surface ($j = M + 1$), where M is the number of vertical layers of the atmosphere

G and λ_{tot} have to be identical for a given climate model regardless of the method used to diagnose the feedback processes. In general, the relation between the sensitivity factor of the climate system and feedback gains in the system can be symbolically written as

$$G = G_0 f(g_1, g_2, \ldots, g_N) \tag{22}$$

where G_0 is the initial sensitivity factor measuring the surface temperature sensitivity without any feedbacks, g_i , a dimensionless factor, is the feedback gain of a feedback process in the climate system, $f(g_1, g_2,...,g_N)$, a dimensionless function of $\{g_i\}$, is the ratio of the (total) climate sensitivity factor to the initial climate sensitivity factor, which measures the net gain with respect to the initial sensitivity factor by all feedbacks in the climate system. When $f(g_1, g_2,...,g_N) > 1$, the net effect of the feedbacks amplifies the initial climate sensitivity. Conversely, when $f(g_1, g_2,...,g_N) < 1$, the feedbacks reduce the climate sensitivity.

As summarized in Table 1, the difference between the PRP and TFRAM roots from the differences in the definition of the feedback gains $\{g_i\}$ and the form of $f(g_1, g_2,...,g_N)$. In the PRP, a feedback gain (denoted as g_i) is measured by the ratio of the TOA radiative perturbation, due to the change in the feedback agent in response to one unit of the surface temperature change, to the TOA radiative perturbation due to one unit change of the atmosphere–surface column temperature. In contrast, a feedback gain (denoted as $\tilde{g}_i$) in the TFRAM is defined as the ratio of the TOA radiative perturbation, due to the change in feedback agent, to the TOA radiative perturbation due to the external forcing.

The difference in the definition of the feedback gains implies that the function $f(g_1, g_2,...,g_N)$ in the PRP is equal to $(1 - \sum_i g_i)^{-1}$ and in the TFRAM is equal to $(1 + \sum_i \tilde{g}_i)$. As in the PRP, the feedback gains are additive (i.e., $\sum_i g_i$ is the total feedback gain in the PRP and $\sum_i \tilde{g}_i$ total feedback gain in the TFRAM). Also as in PRP, a positive $\tilde{g}_i$ means a positive feedback and a negative $\tilde{g}_i$ a negative feedback. In the TFRAM, the additive property of feedback gains automatically implies that the effects of feedbacks are also additive. However, in the PRP, only feedback gains, not the effects of feedbacks, are additive. The difference between the PRP and TFRAM really reflects two ways of solving the linear equation (16) as discussed above.

The difference between the PRP and TFRAM feedback gains reflects the difference of the feedback definition between the PRP and TFRAM methods. In the PRP, the changes in feedback agents are considered as the their responses to the surface temperature change. Physically speaking, there is no reason to put the special priority to the surface temperature change, and to scale the other aspects (feedback agents) of the system response with respect to the surface temperature change as in the PRP. In the TFRAM (or CFRAM), the changes in all climate variables are considered as the system responses to the external forcing, and the change in the surface temperature (or atmosphere–surface column temperature) is only part of the system responses. The other changes are regarded as "feedback agents" to temperature changes only because they contribute to energy flux perturbations that act to enhance or weaken the energy flux perturbation due to the external forcing.

The biggest advantage of the TFRAM gain $\tilde{g}_i$ over the PRP gain g_i is that we can directly calculate the (partial) change in the surface temperature solely due to each of the feedbacks subject to the external forcing provided that the TFRAM feedback gain $\tilde{g}_i$ is known, as summarized in Table 1. These (partial) temperature changes are additive and the sum is equal to the total surface temperature change in response to the external forcing. However, one cannot determine the partial contribution to the total surface temperature change by an individual feedback process from its feedback parameter using the PRP method because the effects of feedbacks are not additive in the PRP.

4.2 Feedback gain matrices in the CFRAM

The feedback gains defined in the TFRAM can be naturally extended to the CFRAM. Similarly to (21), the total temperature change in response to the external forcing can be written as

$$\Delta\vec{\mathbf{T}}^{tot} = \mathbf{G}\Delta\vec{\mathbf{F}}^{ext} = \mathbf{G}_0\left(\mathbf{I} + \sum_{n>0}\tilde{\mathbf{g}}^{(n)}\right)\Delta\vec{\mathbf{F}}^{ext} \tag{23}$$

where $\mathbf{G}$ is the climate sensitivity matrix, $\mathbf{G}_0$ is the initial climate sensitivity matrix corresponding to the climate sensitivity matrix without any feedbacks, $\mathbf{I}$ is the unit matrix, and $\tilde{\mathbf{g}}^{(n)}$ is a feedback gain matrix of a particular feedback process in the climate system. Substituting (9) and (10) into (23), we obtain that

$$\mathbf{G}_0 = \left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1} = \begin{pmatrix} r_{1,1} & \cdots & r_{1,M+1} \\ \vdots & \ddots & \vdots \\ r_{M+1,1} & \cdots & r_{M+1,M+1} \end{pmatrix} \tag{24}$$

$$\tilde{\mathbf{g}}^{(n)} = \begin{pmatrix} \tilde{g}^{(n)}_{1,1} & 0 & \cdots & 0 \\ 0 & \tilde{g}^{(n)}_{2,2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \tilde{g}^{(n)}_{M+1,M+1} \end{pmatrix} \tag{25}$$

$$\tilde{g}^{(n)}_{i,i} = \frac{\sum_{m=1}^{M+1} r_{i,m}\Delta F^{(n)}_m}{\sum_{m=1}^{M+1} r_{i,m}\Delta F^{ext}_m} \quad \text{for} \quad i = 1, 2, \ldots, M+1 \quad \text{and}$$
$$g^{(n)}_{i,j} \equiv 0 \quad \text{for} \quad i \neq j \tag{26}$$

It is rather straightforward to show that the feedback gain matrices are additive and the effects (or temperature changes due to individual feedbacks) are also additive. Also a positive $g_{i,i}$ strengthens the (direct) response to the external forcing alone at the layer i, implying a positive feedback and a negative $g_{i,i}$ means a negative feedback. Based on (26), the elements of a feedback gain matrix are the ratio of the partial temperature change at each layer due to the change in the feedback agent under consideration to the partial temperature at the same layer due to the external forcing alone.

5 Contributions to the lapse rate feedback from individual processes

As shown above, the lapse rate feedback only exists in the TOA-based climate feedback analysis framework (PRP and TFRAM). In this section, we wish to illustrate that the lapse rate feedback really reflects the collective effect of all individual physical and dynamical feedback processes from the prospective of the CFRAM. From the physics point of view, the consideration of a uniform temperature change due to the external forcing or due to an individual feedback in a TOA-only approach is equivalent to considering a climate system response without an atmosphere. Because of the presence of radiative flux perturbations in the atmosphere and changes in the non-radiative exchange of energy between the atmosphere and surface, the energy flux perturbations at the TOA (either due to an external forcing or feedbacks) are not the same as those at the surface. As a result, the uniform change cannot account for the actual surface temperature change because the latter is really determined by the energy flux perturbations at the surface. To make this point more clear, we decompose the lapse rate feedback parameter defined in (18) for the PRP and the partial temperature change due to lapse rate feedback defined in (20) for the TFRAM according to individual physical and dynamical processes of a climate system. It is straightforward to show that

$$\begin{aligned} &\sum_{j=1}^{M+1}\left(-\frac{\partial R^{toa}}{\partial T_j}\right)(\Delta T_j - \Delta T_s) \\ &= \sum_n\left\{\sum_{j=1}^{M+1}\left(-\frac{\partial R^{toa}}{\partial T_j}\right)\left(\Delta T^{(n)}_j - \Delta T^{(n)}_{M+1}\right)\right\} \\ &= \sum_n\left\{\sum_{j=1}^{M+1}\left(-\lambda_p r_{M+1,j}\Delta F^{(n)}_j\right) - \sum_{j=1}^{M+1}\Delta F^{(n)}_j\right\} \end{aligned} \tag{27}$$

where, λ_p is the Planck feedback parameter defined in (18); $\{\Delta T^{(n)}_j,\ j = 1,\ 2,\ldots,\ M+1\}$, with $j = M+1$ corresponding to the surface layer in the model, is the vertical profiles of the partial temperature changes due to the energy flux perturbation $\{\Delta F^{(n)}_j\}$ calculated using (9); $\{r_{M+1,j}\}$ are the element at the jth column in the last row (representing the surface layer) of the matrix $\left(\frac{\partial\vec{\mathbf{R}}}{\partial\vec{\mathbf{T}}}\right)^{-1}$ in (24); $\sum_{j=1}^{M+1}\Delta F^{(n)}_j$ represents the vertically integrated energy flux perturbation due to a specific feedback process "n". Specifically, $\sum_{j=1}^{M+1}\Delta F^{(n)}_j$ is equal to the net downward radiative flux perturbation at the TOA for an external forcing or a physical feedback process (e.g., water vapor, cloud, and surface albedo feedback); $\sum_{j=1}^{M+1}\Delta F^{(n)}_j$ is zero for the local dynamical feedback process that redistributes energy vertically (e.g., sensible and latent heat fluxes and convections); and $\sum_{j=1}^{M+1} \Delta F_j^{(n)}$ is equal to the net horizontal energy transport into the atmosphere–surface column for a non-local dynamical feedback process. In deriving (27), we have first made use of (10), then summed up all rows of (9) from $j = 1$ to $j = M + 1$ after multiplying both sides of (9) with $\left(\frac{\partial \vec{\mathbf{R}}}{\partial \mathbf{T}}\right)$, and next used $\Delta T_{M+1}^{(n)} = \sum_{j=1}^{M+1} (r_{M+1,j} \Delta F_j^{(n)})$ according to (9).

We note that the term $r_{M+1,\ j}\ \Delta F_j^{(n)}$ for $j < (M + 1)$ represents the surface temperature change due to the "back radiation" effect of the air temperature change in response to the feedback energy flux perturbation $\Delta F_j^{(n)}$ at the jth layer and $\sum_{j=1}^{M} (r_{M+1,j} \Delta F_j^{(n)})$ is the total "back radiation" effect of the atmosphere temperature change. The term $r_{M+1,M+1}\ \Delta F_{M+1}^{(n)}$ is the direct response of the surface temperature to the energy flux perturbation at the surface layer $\Delta F_{M+1}^{(n)}$. Then, $\sum_{j=1}^{M+1} (-\lambda_p r_{M+1,j} \Delta F_j^{(n)})$ would be the TOA radiative energy perturbation caused by a uniform change of the system temperature equal to $\Delta T_{M+1}^{(n)}$ in response to the feedback energy flux perturbations at all layers.

Dividing (27) by ΔT_s (ΔT_s is the total surface temperature change) yields the lapse rate feedback parameter defined in (18) for the PRP, and dividing (27) by $(-\lambda_p)$ gives rise to the partial temperature change due to the lapse rate feedback in (20) for the TFRAM. Based on (27), the contribution to the lapse rate feedback from a feedback process "n" can be regarded as the TOA energy flux perturbation caused by the residual between the actual surface temperature response to height-dependent energy flux perturbation of the feedback "n" and the hypothetical uniform response to the TOA part of the energy perturbation. Therefore, the lapse rate feedback is equal to the sum of the residuals between the actual responses and uniform responses due to the external forcing alone, and due to each of individual feedbacks. For a feedback process that has a non-zero energy flux perturbation at the TOA ($\sum_{j=1}^{M+1} \Delta F_j^{(n)} \neq 0$), only part of its feedback effect, equal to difference between its (total) non-uniform response and uniform response, is included in the lapse rate feedback. For a local dynamical feedback process that has no contribution to the energy flux perturbation at the TOA ($\sum_{j=1}^{M+1} \Delta F_j^{(n)} = 0$), all of its effect is lumped into the lapse rate feedback.

6 Summary and discussions

This paper presents the mathematical formulation of a coupled atmosphere–surface climate feedback–response analysis method, referred to as the CFRAM, for analyzing climate feedback and sensitivity in coupled general circulation models with a full set of physical parameterization packages. The CFRAM is based on the consideration that climate change is not only determined by the (radiative) energy exchange between the climate system and outer space, but also intrinsically constrained by the energy "flow" within the climate system. In other word, in response to an external forcing that has a vertical and horizontal structure, both radiative and non-radiative energy flux perturbations can be regarded as feedbacks because they do act to either strengthen, or weaken, or even oppose the external forcing. Accordingly, we consider the conservation equation of total energy for atmosphere–surface (land/ocean) column in formulating the CFRAM. We consider that all of these radiative and non-radiative feedbacks result from the climate system response to the external forcing, rather than solely from the surface temperature change (directly or indirectly).

This generalized concept of feedback–response enables us to apply the energy balance equation in a different way from the traditional climate feedback analysis framework. In this new framework, we take advantage of the fact that the infrared radiation is explicitly and directly related to the atmosphere–surface temperatures. Then we can infer the (total) temperature change by requiring its infrared radiation perturbation to exactly balance the external forcing and feedbacks (including all non-temperature induced radiative energy flux perturbations and non-radiative energy flux perturbations). Specifically, with the CFRAM as an offline diagnostic tool, we can calculate the partial temperature change associated with an individual energy flux perturbation by requiring the infrared radiation induced by the temperature change alone to exactly balance the energy flux perturbation under consideration. The resultant partial temperature changes are addable and their sum gives rise to the total temperature response to the external forcing. Therefore, the total temperature change calculated from the CFRAM can be directly compared with its counterpart produced in the CGCM climate model simulations. This enables us to estimate the accuracy/uncertainty of the temperature changes due to individual feedbacks calculated in the CFRAM. In this sense, we "isolate" the contributions to the (total) temperature change from the external forcing alone, and from individual feedbacks although physically speaking, these feedbacks are not independent with one another.

The definitions of climate forcing, feedback, and response adopted in a TOA-based approach (e.g., the PRP, or the TFRAM, the TOA version of the CFRAM) can be regarded as the special case of the generalized definitions of climate forcing, feedback, and response adopted in the CFRAM. The vertical integration of the generalized climate forcing from the surface to the TOA is exactly equal to the TOA-based climate forcing. In terms of climate feedback, the surface-to-TOA vertical integrations of the water vapor feedback, cloud feedback and albedo feedback based on the new definitions are also identical to their TOA-based counterparts. Also, the surface-to-TOA vertical integration of the non-local dynamic feedback due to changes in horizontal energy redistribution processes (e.g., poleward heat transport) is identical to its TOA-based counterpart. However, in a TOA-based approach, the climate response is measured only in terms of surface temperature. Therefore, the TOA radiative perturbation due to the difference between the surface and atmosphere temperature changes is also regarded as a climate feedback (referred as "lapse-rate" feedback) in a TOA-based approach (including both the PRP and the TOA version of the CFRAM or the TFRAM). In the CFRAM, the climate response is defined as the (simultaneous) changes in both the surface and atmosphere (including its vertical profile). As a result, the lapse-rate feedback is no longer a feedback in the CFRAM. In the CFRAM, the energy flux perturbations induced by the changes in the thermodynamical and dynamic processes that redistribute energy vertically but cause no net radiative energy fluxes at the TOA are also feedbacks (e.g., vertical convections and surface sensible heat and latent heat fluxes). In a TOA based approach, the effects of these vertical energy redistribution processes are implicitly included in the lapse rate feedbacks. Furthermore, the lapse rate feedback defined in a TOA-based approach also includes the vertically non-uniform effects of the climate forcing, water vapor feedback, cloud feedback, surface albedo feedback, and non-local dynamic feedbacks associated with changes in horizontal energy transport processes because they all contribute to the difference between the surface and atmosphere temperature changes.

In summary, the unique features of the coupled atmosphere–surface CFRAM in comparison with the existing feedback analysis methods include:

- The surface temperature change in response to an individual feedback process in the coupled atmosphere–surface CFRAM includes both the direct change by the feedback process and the indirect change by the "back radiation" due to the change in the air temperature caused by the same feedback process. Therefore, the (air) temperature feedback, by definition, no longer exists in the coupled atmosphere–surface CFRAM.
- In the CFRAM, the feedback decomposition is based on the thermodynamic and dynamical processes that directly affect either the energy exchange between the climate system and outer space, or the internal energy flow in the system. The energy flux perturbations due to the changes in these processes can be directly inferred from individual parameterization packages and dynamical modules. The process-based feedback decomposition adopted in the CFRAM allows us to examine not only those feedbacks that directly affect the TOA radiative fluxes, such as water vapor, clouds, and ice-albedo feedbacks, but also those feedbacks that do not directly affect the TOA radiation, such as evaporation, convections, and horizontal heat transport, in a unified framework.
- In the CFRAM, the feedback gain matrices measure the strength of individual feedbacks. The feedback gain matrices can be estimated from the energy flux perturbations inferred from individual parameterization packages and dynamical modules. The inter-model spread of a feedback gain matrix would help us to detect the origins of the uncertainty of future climate projections in climate model simulations.
- The CFRAM can be easily used in various configurations, such as the global mean response with the vertical profile (1D), or local/regional responses (e.g., low latitude versus high latitudes or ocean versus land), or the zonally averaged response with the meridional and vertical profile (2D), or even in a fully 3D configuration. In Part II of this two-part series paper, we illustrate the CFRAM in the context of a coupled atmosphere–surface single column climate model and compare it with the other feedback analysis methods.

Finally, we wish to comment on whether it is possible to further reduce the uncertainty in estimating the climate sensitivity using GCMs, an issue recently raised by Roe and Baker (2007). Based on the definition of feedback gains in the PRP, Roe and Baker (2007) pointed out that the uncertainty in estimating the climate sensitivity is proportional to the product of the square of the total gain of the climate system and uncertainty in individual feedback gains. Because the square of the total gain of the climate system is in the order of 10, a small uncertainty in individual feedback gains would get amplified, leading to a very large uncertainty in estimating the climate sensitivity. This prompts them to conclude that there would be little possibility to further reduce the uncertainty in estimating the climate sensitivity even though uncertainties of individual feedbacks could be reduced significantly. However, if one uses the feedback gains defined in the TOA version of the CFRAM (TFRAM), one would find that the uncertainty in estimating the climate sensitivity is linearly proportional to uncertainties in individual feedback gains and is not related to the total gain at all. This is because in the TFRAM, the climate sensitivity is linearly proportional to the sum of all feedback gains. Since the total gain or the climate sensitivity defined in the PRP and TFRAM has to be identical for a given climate system, it is believed that the uncertainty in estimating the climate sensitivity can be further reduced if uncertainties in estimating individual feedbacks can be reduced.

Acknowledgments The authors are grateful for the constructive comments from three anonymous reviewers. This work is supported by grants from the NOAA/Office of Global Programs (GC04-163 and GC06-038).

References

Bates JR (2007) Some considerations of the concept of climate feedback. Q J R Meteorol Soc 133:545–560

Bode H (1945) Network analysis and feedback amplifier design. Van Nostrand, pp 551

Boer GJ, Yu B (2003) Climate sensitivity and response. Clim Dyn 20:415–429

Bony S et al (2006) How well do we understand and evaluate climate feedback processes? J Clim 19:3445–3482

Cai M (2006) Dynamical greenhouse-plus feedback and polar warming amplification. Part I: A dry radiative–transportive climate model. Clim Dyn 26:661–675

Cai M, J-H Lu (2007) Dynamical greenhouse-plus feedback and polar warming amplification. Part II: Meridional and vertical asymmetries of the global warming. Clim Dyn 29:375–391

Cess RD et al (1990) Intercomparison and interpretation of cloud–climate feedback processes in nineteen atmospheric general circulation models. J Geophys Res 95:16601–16615

Colman R (2003) A comparison of climate feedbacks in general circulation models. Clim Dyn 20:865–873

Fu Q, Liou KN (1993) Parameterization of the radiative properties of cirrus clouds. J Atmos Sci 50:2008–2025

Hall A, Manabe S (1999) The role of water vapour feedback in unperturbed climate variability and global warming. J Clim 12: 2327–2346

Hall A, Qu X (2006) Using the current seasonal cycle to constrain snow albedo feedback in future climate change. Geophys Res Lett 33:L03502. doi:10.1029/2005GL025127

Hansen J et al (1984) Climate sensitivity: analysis of feedback mechanisms. In: Climate processes and climate sensitivity. Geophysics monography, vol 29. American Geophysics Union, pp 130–163

Held IM, Soden BJ (2000) Water vapor feedback and global warming. Annu Rev Energy Environ 25:441–475

Peixoto JP, Oort AH (1992) Physics of climate. American Institute of Physics, pp 308–364

Ramaswamy V et al (2001) Radiative forcing of climate change. In: Houghton JT et al (eds) Climate change 2001: the scientific basis. Cambridge University Press, Cambridge, pp 349–416

Roe GH, Baker MB (2007) Why is climate sensitivity so unpredictable? Science 318:629–632

Schneider EK, Kirtman BP, Lindzen RS (1999) Tropospheric water vapor and climate sensitivity. J Atmos Sci 56:1649–1658

Soden BJ, Held IM (2006) An assessment of climate feedbacks in coupled ocean atmosphere models. J Clim 19:3354–3360

Soden BJ, Broccoli AJ, Hemler RS (2004) On the use of cloud forcing to estimate cloud feedback. J Clim 17:3661–3665

Stephens GL (2005) Cloud feedbacks in the climate system: a critical review. J Clim 18:237–273

Wetherald R, Manabe S (1988) Cloud feedback processes in a general circulation model. J Atmos Sci 45:1397–1415

Winton M (2006) Surface albedo feedback estimates for the AR4 climate models. J Clim 19:359–365

Zhang MH, Hack JJ, Kiehland JT, and Cess CD (1994) Diagnostic study of climate feedback processes in atmospheric general circulation models. J Geophys Res 99:5525–5537