Transient climate response in a two-box energy-balance model. Part I: analytical solution and parameter calibration using CMIP5 AOGCM experiments.

Olivier Geoffroy * and David Saint-Martin

Centre National de Recherches Météorologiques (CNRM-GAME), Toulouse, France

Dirk Jan Leo Olivié

Center for International Climate and Environmental Research - Oslo (CICERO), Oslo, Norway

and

University of Oslo, Oslo, Norway

Aurore Voldoire, Gilles Bellon and Sophie Tytéca

Centre National de Recherches Météorologiques (CNRM-GAME), Toulouse, France

* *Corresponding author address:* Olivier Geoffroy, Centre National de Recherches Météorologiques (CNRM-GAME) 42 av. G. Coriolis, 31057 Toulouse, France.

E-mail: olivier.geoffroy@meteo.fr

ABSTRACT

This is the first part of a series of two articles analyzing the global thermal properties of atmosphere-ocean coupled General Circulation Models (AOGCMs) within the framework of a two-box Energy Balance Model (EBM). In this part, the general analytical solution of the system is given and two idealized climate change scenarios, one with a step forcing and one with a linear forcing, are discussed. These solutions give a didactic description of the contributions from the balanced response, and from the fast and slow transient responses during a climate transition. Based on these analytical solutions, we introduce a simple and physically-based procedure to calibrate the two-box model parameters using an AOGCM step-forcing experiment. Using this procedure, the global thermal properties of twelve AOGCMs participating in CMIP5 are determined. It is shown that, for a given AOGCM, the EBM tuned with only the abrupt 4xCO_2 experiment is able to reproduce with a very good accuracy the temperature evolution in both a step-forcing and a linear-forcing experiments. The role of the upper-ocean and the deep-ocean heat uptakes in the fast and slow responses is also discussed. One of the main weakness of the simple EBM discussed in this part is its ability to represent the evolution of the top-of-the-atmosphere radiative imbalance in the transient regime. This issue is addressed in Part II by taking into account the efficacy factor of deep-ocean heat uptake.

1. Introduction

Determining the response of the climate system to an imposed external perturbation is a major challenge in climate science. The global and annual mean surface temperature response is a useful metric to determine the magnitude of a climate change induced by an externaly imposed radiative perturbation. Indeed, many studies suggest that most of the climate variables are related to the global mean surface temperature response. Coupled atmosphere-ocean general circulation models (AOGCMs) are the most comprehensive tool to study climate changes and perform climate projections. They can be used to assess the changes in global temperature but they are computationally expensive. Alternatively, simple climate models (SCMs), which estimate approximately the global mean surface temperature change for a given, externally-imposed perturbation in the Earth's radiation balance (Meinshausen et al. 2008; Good et al. 2011), can be used to emulate the AOGCM responses in order to cover a wide range of scenarios with a negligible computational cost.

Energy-balance models (EBMs) are physically-based SCMs. They are useful to summarize AOGCM global thermal properties, intercompare and analyze AOGCM responses (Raper et al. 2002; Soden and Held 2006; Gregory and Forster 2008; Dufresne and Bony 2008). In the case of a small perturbation, some EBMs assume that the thermal energy balance of the climate system is expressed as a linear function of temperature perturbation only (Budyko 1969; Sellers 1969). The net radiative imbalance due to an external forcing and a temperature change can be expressed as $N = \mathcal{F} - \lambda T$. The radiative feedback parameter λ with respect to the global mean surface air temperature T depends on the type of forcing (Hansen et al. 2005). The imposed radiative forcing $\mathcal{F}$ includes the effects of both fast (few months) stratospheric and tropospheric adjustments (Gregory and Webb 2008). In this formulation of the radiative imbalance N, the assumption of linear dependency in T suffers from some limitations (Gregory et al. 2004; Williams et al. 2008; Winton et al. 2010; Held et al. 2010) that are adressed in Part II of this study.

In equilibrium, $N = 0$ and the steady-state temperature is equal to $T_{eq} = \mathcal{F}/\lambda$. The equilibrium climate sensitivity (ECS), which is defined as the equilibrium mean surface air temperature perturbation resulting from a doubling carbon dioxide radiative forcing, is commonly used as a metric of anthropogenic climate change. However, this metric is not sufficient to study the transient regime because of the climate-system thermal inertia. Indeed, the rate of change in the heat content of the climate system is equal to the Earth's radiative imbalance and this change occurs on large timescales due to the large thermal inertia of the deep ocean (Dickinson 1981; Hasselmann et al. 1993; Murphy 1995; Gregory 2000; Held et al. 2010). Based on empirical relationships, Gregory and Mitchell (1997) and Raper et al. (2002) propose a formulation for the deep-ocean heat uptake proportional to the surface temperature perturbation: $H = \kappa T$. However, this formulation is not able to represent the equilibrium temperature response in the case of a step-forcing or a stabilization scenario because the deep-ocean temperature response is by definition neglected.

The solution to circumvent this shortcoming is to introduce a second layer which represents the deep ocean. Splitting of the climate system into two thermal reservoirs with different heat capacities allows one to take into account the ocean thermal saturation along a transient regime until equilibrium and then to represent the two distinct timescales to the global mean climate system response (Hasselmann et al. 1993; Held et al. 2010). This sytem is similar to the three-box EBM presented in Dickinson (1981), the atmosphere and the upper-ocean layers being considered as one single layer characterized by the surface air temperature.

In this study, we analytically derive the solution of this two-box energy-balance model and propose a calibration method for determining the equivalent thermal parameters of a given AOGCM. We then assess the validity of this simple framework to represent the behavior of the complex coupled models in response to an idealized forcing scenario by analyzing the results of twelve AOGCMs participating in the fifth phase of the Coupled Model Intercomparison Project (CMIP5). The role of each box heat uptake in the fast and slow components of the transient response is also discussed.

The structure of the paper is as follows: after introducing the theorical framework and describing the analytical solutions for different forcing scenarios in Section 2, the methodology used to adjust the two-box EBM response to AOGCMs results is presented and applied to CMIP5 AOGCMs in Section 3.

2. Theoretical framework

a. Two-box energy-balance model

We consider the linear two-box energy-balance model described in Held et al. (2010). Held et al. (2010) also proposed an alternative model with an additional parameter, an efficacy factor for deep-ocean heat uptake that will be discussed in Part II. The climate system is split in two layers (Gregory and Mitchell 1997; Gregory 2000). The first one corresponds to the atmosphere, the land surface and the upper ocean, and the second one represents the deep ocean. The state of each layer is described by a temperature perturbation T and T_0. T is usually taken as the global mean surface air temperature perturbation from the control climate. T_0 is a characteristic temperature perturbation of the deep ocean. T and T_0 verify the following system of equations:

$$C\frac{dT}{dt} = \mathcal{F} - \lambda T - \gamma(T - T_0), \tag{1}$$

$$C_0\frac{dT_0}{dt} = \gamma(T - T_0). \tag{2}$$

This system has two prognostic variables and five free parameters: λ, γ, C, C_0 and a radiative forcing amplitude parameter. Whatever the radiative forcing agent, the radiative forcing formulation requires at least one model-dependant reference radiative parameter $\mathcal{F}_{ref}$ due to stratospheric and tropospheric adjustments. In the case of a CO_2 perturbation, the radiative forcing can be expressed as a function of the CO_2 concentration and a radiative parameter following (IPCC 1990):

$$\mathcal{F}(t) = \frac{\mathcal{F}_{2\mathrm{xCO}_2}}{\ln(2)} \ \ln\left(\frac{[\mathrm{CO}_2]_t}{[\mathrm{CO}_2]_0}\right). \tag{3}$$

where $[\mathrm{CO}_2]_t$ is the time-dependant carbon dioxide concentration, $[\mathrm{CO}_2]_0$ is the preindustrial CO_2 concentration and $\mathcal{F}_{2\mathrm{xCO}_2}$ is the net radiative forcing associated with a doubling of the atmospheric CO_2 concentration.

$C\frac{dT}{dt}$ and $C_0\frac{dT_0}{dt}$ are the tendencies of heat contents respectively of the upper and the lower layer. C and C_0 are effective surfacic heat capacities respectively of the upper (by neglecting atmosphere and land surface heat capacities) and the deep ocean. The parameter γ is a heat exchange coefficient. The heat flux exchange between the two layers is thus assumed to be proportional to the difference between the two temperature perturbations. In the limit of an infinite deep-ocean heat capacity ($C_0 \to \infty$), T_0 is zero and the expression of the heat flux exchange is the one proposed by Gregory and Mitchell (1997) with $\kappa = \gamma$. In this one-box model (the deep-ocean layer is an external infinite reservoir), the temperature perturbation verifies the following equation (Raper et al. 2002; Dufresne and Bony 2008):

$$C\frac{dT}{dt} = \mathcal{F} - \lambda T - \kappa T. \tag{4}$$

The differences between the two models are analyzed in the next section.

The temperature T_H associated with the climate system heat-uptake is defined as the disequilibrium temperature difference between T and the instantaneous equilibrium temperature $T_{eq}(t) = \mathcal{F}(t)/\lambda$ (Winton et al. 2010). The latter is the equilibrium temperature associated with the instantaneous forcing applied at time t. The heat-uptake temperature represents the instantaneous rate of heat storage in the climate system:

$$T_H(t) = T(t) - T_{eq}(t) = -\frac{1}{\lambda}\left[C\frac{dT}{dt} + C_0\frac{dT_0}{dt}\right]. \tag{5}$$

Contrary to Winton et al. (2010), a negative heat-uptake temperature corresponds here to a positive heat storage in the climate system.

b. Analogy with electricity

The two-box energy-balance model and its simpler version (one-box model) can be advantageously described in terms of equivalent electrical circuits (Fig. 1). While temperature differences are analogous to electrical potential differences, heat fluxes are analogous to intensities.

In the case of the one-box model (see Fig. 1a), the first layer is a capacitor with capacity C. It is linked to the external system by a resistance $1/\lambda$ and to the second layer by a resistance $1/\kappa$. The input voltage is equal to the instantaneous equilibrium temperature $T_{eq}(t) = \mathcal{F}(t)/\lambda$ and the output voltage is the surface mean temperature T. The intensity in the main branch of the circuit is the radiative imbalance $N = \mathcal{F} - \lambda T$. For a step-forcing, the capacitor voltage increases until saturation. The intensity through the capacitor becomes zero and the equilibrium temperature response is given by a voltage divider and is equal to $\mathcal{F}/(\lambda + \kappa)$.

In the case of the two-box model (see Fig. 1b), there is a resistance $1/\gamma$ and an additional capacitor with a higher capacity value C_0 in the secondary branch through which the intensity analogous to the deep-ocean heat uptake flows. The deep-ocean temperature perturbation T_0 is the voltage across this capacitor. In equilibrium, both intensities are zero and $T = T_0 = T_{eq}$.

Both circuits are low-pass filters. The Bode diagram of the second one is given in Appendix C. It is interesting to note that in the framework of electrical circuits, the forcing is directly seen as an input perturbation in temperature T_{eq} instead of a perturbation in radiative flux, from which the output temperature T can be derived by applying a transfer function $\mathcal{H}$. Indeed, these functions are apparent in the analytical solutions that are given in the following section.

c. *Analytical solutions*

In Appendix A, the general solutions of both the mean surface temperature and the deep-ocean temperature responses are derived for any forcing function $t \to \mathcal{F}(t)$. With an integration by parts of the equations (A8) and (A9), the temperature perturbations T and T_0 can be written as the sum of the balanced temperature $T_{eq}(t)$ and two modes characterized by two distinct timescales, τ_f (fast) and τ_s (slow):

$$T(t) = T_{eq}(t) - \sum_{i=\{s,f\}} \frac{a_i}{\lambda} \left(\mathcal{F}(0) e^{-t/\tau_i} + \int_0^t \mathcal{F}'(\xi) e^{-(t-\xi)/\tau_i} d\xi \right), \tag{6}$$

$$T_0(t) = T_{eq}(t) - \sum_{i=\{s,f\}} \frac{\phi_i a_i}{\lambda} \left(\mathcal{F}(0) e^{-t/\tau_i} + \int_0^t \mathcal{F}'(\xi) e^{-(t-\xi)/\tau_i} d\xi \right), \tag{7}$$

where τ_i, a_i and ϕ_i are parameters depending on C, C_0, γ and λ. Their expressions are given in Table 1. Note, in particular, that $a_f + a_s = 1$ and $\phi_f a_f + \phi_s a_s = 1$, and that $\phi_f < 0$ while a_f, a_s, and ϕ_s are all positive.

The sum term in Eq. (6) is the heat-uptake temperature T_H. The latter is the sum of two modes that can be decomposed in two terms depending on the forcing function. The first contribution is an instantaneous deviation associated to a discontinuity of the forcing at $t = 0$. The second one is due to the time-evolution of the forcing.

In the following paragraphs, we briefly discuss the analytical solution for two idealized forcings: step and linear. In Appendices B and C, we present solutions for stabilization, abrupt return to zero and periodic forcings.

1) Step forcing

For a step forcing:

$$\mathcal{F}(t) = \begin{cases} 0 & \text{if } t < 0 \\ \mathcal{F} & \text{if } t \geq 0, \end{cases} \tag{8}$$

the analytical solution of the two-box energy-balance model is given by:

$$T(t) = \frac{\mathcal{F}}{\lambda}\left[a_f(1-e^{-t/\tau_f}) + a_s(1-e^{-t/\tau_s})\right], \tag{9}$$

$$T_0(t) = \frac{\mathcal{F}}{\lambda}\left[\phi_f a_f(1-e^{-t/\tau_f}) + \phi_s a_s(1-e^{-t/\tau_s})\right]. \tag{10}$$

By decomposing the response as the sum of the equilibrium temperature response and the two modes characterized by the two distinct timescales, τ_f (fast) and τ_s (slow) following Eqs. (6) and (7), the temperature perturbations T and T_0 are :

$$T(t) = T_{eq} - a_f T_{eq} e^{-t/\tau_f} - a_s T_{eq} e^{-t/\tau_s}, \tag{11}$$

$$T_0(t) = T_{eq} - \phi_f a_f T_{eq} e^{-t/\tau_f} - \phi_s a_s T_{eq} e^{-t/\tau_s}. \tag{12}$$

Thus, a_i is the partial contribution of the mode i to the T_H initial amplitude in the case of a step forcing. Initially, both the slow and the fast terms are negative with respective amplitudes $-a_f T_{eq}$ and $-a_s T_{eq}$. During the transition, they increase exponentially towards zero with their respective relaxation times τ_f and τ_s as illustrated in Fig. 2a, b.

In Eqs. (11) and (12), we can also see that the last two terms are the projections of the perturbations $T_H(t) = T(t) - T_{eq}$ and $T_{0H}(t) = T_0(t) - T_{eq}$ from the new equilibrium $T(t) = T_0(t) = T_{eq}$ onto the eigenmodes of the linear system of equations. Since $\phi_s > 0$, the projection of T_H and T_{0H} onto the slow eigenmode have the same sign, and the slow eigenmode corresponds to a joint adjustment of the upper and lower layers. On the other hand, since $\phi_f < 0$, the projection of T_{0H} onto the fast eigenmode is of opposite sign to the projection of T_H (in the fast mode, $T_H < 0$ and $T_{0H} > 0$). The perturbation heat flux from the lower layer to the upper layer is $-H = -\gamma(T_H - T_{0H})$ and its projection onto the fast eigenmode is of opposite sign to T_H. The fast eigenmode thus corresponds to an adjustment of the upper layer by both the radiation imbalance and the deep-ocean heat uptake. The two physical processes at play interact positively to adjust the smallest energy reservoir. This explains why the characteristic timescale τ_f is shorter than the characteristic timescale of a one-box model of the upper layer without deep-ocean heat uptake, that is the limit of τ_s

when C_0 tends toward zero: $\tau_f < C/\lambda$. Still, τ_f is longer than the characteristic timescale of the one-box model of the upper layer with deep-ocean heat uptake presented in Gregory and Mitchell (1997) and Raper et al. (2002), that is the limit of τ_f when C_0 tends toward infinity: $\tau_f > C/(\gamma + \lambda)$. In that model, the deep-ocean heat uptake damps T_H more efficiently than in the two-box model because of its infinite heat capacity.

2) Linear forcing

To derive the analytical solution of the system for a linear forcing:

$$\mathcal{F}(t) = \begin{cases} 0 & \text{if } t < 0 \\ F\,t & \text{if } t \geq 0, \end{cases} \tag{13}$$

we have to compute the integral $\mathcal{I}(t) = \int_0^t \xi e^{\xi/\tau_i} d\xi$. As we found $\mathcal{I}(t) = \tau_i t e^{t/\tau_i} + \tau_i^2(1 - e^{t/\tau_i})$, the general solution can be written as

$$T(t) = \frac{F}{\lambda}t - \frac{F}{\lambda}\tau_f a_f(1 - e^{-t/\tau_f}) - \frac{F}{\lambda}\tau_s a_s(1 - e^{-t/\tau_s}), \tag{14}$$

$$T_0(t) = \frac{F}{\lambda}t - \frac{F}{\lambda}\phi_f\tau_f a_f(1 - e^{-t/\tau_f}) - \frac{F}{\lambda}\phi_s\tau_s a_s(1 - e^{-t/\tau_s}), \tag{15}$$

As in the step-forcing case, the surface temperature perturbation is the sum of a balanced response $T_{eq}(t) = \mathcal{F}(t)/\lambda$ and an imbalance term which can be decomposed into a fast and a slow mode response as illustrated in Fig. 2c, d. Contrary to the abrupt case, the system is initially in equilibrium and deviates from its instantaneous balanced temperature $T_{eq}(t)$ afterwards. The fast and slow responses decrease with time and asymptotically tend towards negative limits. Their amplitudes are proportional to their respective relaxation times resulting in a negligible amplitude of the fast response.

Assuming a logarithmic relationship between the radiative forcing and the carbon dioxide concentration [Eq. (3)], the 1% y^{-1} CO_2 experiment corresponds to a linear forcing with:

$$F = \frac{\mathcal{F}_{2xCO_2}}{t_{2xCO_2}} \quad \text{with } t_{2xCO_2} \approx 70\ y. \tag{16}$$

3. Multi-model analysis

In this section, a method for tuning the two-box model parameters described above to individual AOGCMs using only the idealized step-forcing experiments is proposed. The tuning method is then applied to twelve available AOGCMs participating in the fifth phase of the Coupled Model Intercomparison Project (Taylor et al. 2011) and is validated by comparison with AOGCM responses to the linear-forcing, 1% y^{-1} CO_2 experiments.

a. Method for parameter calibration

The method uses only an AOGCM non equilibrated response of a step-forcing experiment. We assume that the top of the climate system corresponds to the model top-of-the-atmosphere (TOA). Both radiative net flux change at TOA and surface temperature change T are used to adjust the two radiative parameters $\mathcal{F}_{ref}$ (the adjusted radiative forcing amplitude) and λ. Only T is used to adjust the thermal inertia parameters C, C_0 and γ. The method can thus be decomposed in two steps.

1) 1^{st} STEP

The first step consists in estimating the radiative parameters by using the method described in Gregory et al. (2004). Using this method, the computation of $\mathcal{F}_{ref}$ and λ takes into account both stratospheric and tropospheric adjustments. However it assumes a linear dependancy between the Earth's radiation imbalance and the surface temperature perturbation such that $N = \mathcal{F} - \lambda T$. The limitations of this assumption are discussed in Part II.

2) 2^{nd} STEP

The second step consists first in the calibration of the four mode parameters τ_f, τ_s, a_f and a_s by fitting the global mean surface air temperature response. For $t \gg \tau_f$, Eq. (11) can be approximated as follows:

$$T \approx T_{eq}(1 - a_s e^{-t/\tau_s}) \quad \Rightarrow \quad \log(1 - \frac{T}{T_{eq}}) \approx \log a_s - \frac{1}{\tau_s} t. \tag{17}$$

Assuming that $\tau_f \ll 30\ y$, the linear regression of $\log(T_{eq} - T)$ against t over the period $30 - 150\ y$ gives estimations of τ_s and a_s.

Then $a_f = 1 - a_s$ is known. τ_f can then be expressed from Eq. (11) in function of these three parameters and the surface temperature response:

$$\tau_f = t / \left[\log a_f - \log\left(1 - T/T_{eq} - a_s e^{-t/\tau_s}\right)\right]. \tag{18}$$

Its value is estimated by averaging over the first ten years of the step-forcing experiment.

Finally, the remaining physical parameters of the model (the heat capacities C and C_0 and the heat exchange coefficient γ) are computed from the other parameters using the following analytical relationships (see Table 1):

$$C = \lambda/(a_f/\tau_f + a_s/\tau_s), \tag{19}$$

$$C_0 = \lambda(\tau_f a_f + \tau_s a_s) - C, \tag{20}$$

$$\gamma = C_0/(\tau_f a_s + \tau_s a_f). \tag{21}$$

This methodology is applied to instantaneous carbon dioxide quadrupling (abrupt 4xCO_2) experiments (with a typical time integration of 150 years) performed by an ensemble of twelve AOGCMs participating in the CMIP5.

b. Results

1) Radiative parameters

For the 12 AOGCMs considered here (Table 2), the multimodel average of the net radiative forcing (6.8 W m^{-2}) is very close to previous CMIP3 analysis results (Williams et al. 2008), and the relative intermodel standard deviation is about 14%. The estimates for the model's feedback parameters are consistent with previous results with older AOGCMs (Soden and Held 2006). The multimodel mean (1.11 W m^{-2} K^{-1}) and standard deviation (0.31 W m^{-2} K^{-1}) of the total feedback parameters (Table 2) are close to previous values obtained for CMIP3 models and for different types of scenarios.

The estimation of the 4xCO_2 equilibrium temperature response for each model is presented in Table 2. The equilibrium temperature response ranges from 4.1 K to 9.1 K. The spread among the responses is as large as those of CMIP3 simulations.

2) Climate system inertia parameters

In Table 3, we summarize the corresponding thermal parameters for each of the 12 models. We first note that the deep-ocean heat-capacity values are about an order of magnitude larger than the upper-layer heat-capacity values. The multimodel means of C (7.7 W y m^{-2} K^{-1}) and C_0 (106 W y m^{-2} K^{-1}) are close to Dickinson (1981) estimations of 10 W y m^{-2} K^{-1} and 100 W y m^{-2} K^{-1} for the ocean mixed layer and for the deep-ocean capacities respectively. The deep-ocean heat-capacity mean value is however larger than the Murphy (1995) estimate of the ocean capacity of about 52 W y m^{-2} K^{-1} (1.65 10^9 J m^{-2} K^{-1}). Considering ocean covers $f_0 = 70\%$ of the climate system surface, and using a constant water heat-capacity of $c_p = 4180$ J kg^{-1} K^{-1} and a constant density of salt water $\rho = 1030$ kg m^{-3}, the AOGCMs ensemble mean C_0 value corresponds to an equivalent deep-ocean layer depth D_0 equal to:

$$D_0 = \frac{86400 * 365.25 * C_0}{\rho c_p f_0} = 1110 \text{ m.} \tag{22}$$

Similarly, an upper ocean with an effective surfacic heat capacity equal to the AOGCMs ensemble mean value is equivalent to a 81 m thick mixed layer.

The INM model gives a very large value of C_0 (317 W y m^{-2} K^{-1}) in comparison with other models. One can wonder if this estimation can be biased by the drift in surface temperature evolution since the INM model is one of the two models with the largest drift in surface temperature evolution in the course of the preindustrial control simulation. Indeed, the INM drift is of the order of -0.03 K per century (over a period of 500 y) against a model ensemble mean of absolute value of 0.02 K per century and a standard deviation of 0.012 K per century. However, after removing the temperature trend, the C_0 estimate for INM still remains largely outside the range of the model ensemble with a value of 271 W y m^{-2} K^{-1}. All other parameters of this model and all other model parameters are not significantly impacted by the temperature drift correction. Further investigation would be needed to explain the INM behaviour. By excluding this model, the ensemble mean C_0 value is 87 W y m^{-2} K^{-1} with a much smaller standard deviation of 26 W y m^{-2} K^{-1}.

The heat exchange coefficient γ ranges from 0.5 to 0.9 W m^{-2} K^{-1} with an ensemble mean of 0.7 W m^{-2} K^{-1}. These values are somewhat larger than the one-box EBM heat exchange coefficient κ values estimated by Raper et al. (2002) and Gregory and Forster (2008) and of the same order of magnitude than Plattner et al. (2008) estimates. One could expect that the introduction of the deep-ocean temperature perturbation T_0 in the two-box EBM reduces the contribution of the temperature difference term to the deep-ocean heat uptake $H = \gamma(T - T_0)$ formulation: for a given H, $T - T_0 < T$ so that $\gamma > \kappa$.

Fast and slow time responses are also given in Table 3. The fast time constant is of the order of 4 years and the slow response of the order of 250 years. These values are consistent with previous estimations of climate-system timescales (see Olivié et al. (2011) for example). The intermodel standard deviation for the slow relaxation time is about 150 years. It is reduced to 60 years by omitting the large value of τ_s (due to the large C_0) of the INM model.

The estimates of these climate-system parameters could be biased as a consequence of the biases in the radiative parameters estimated using the method of Gregory et al. (2004). The sensitivity of these estimates to a more refined formulation of the two-box model is explored in Part II.

3) Global mean surface air temperature response

The comparison between the analytical model calibrated from abrupt 4xCO_2 experiment and AOGCM responses to the abrupt 4xCO_2 and the 1% y^{-1} CO_2 increase up to 4xCO_2 is shown in Figs. 3 and 4. For the CNRM and GFDL models, a 2xCO_2 stabilization scenario is also available. Note that the analytical EBM results for the 1% y^{-1} CO_2 and the stabilization cases are computed using the parameters tuned using the abrupt 4xCO_2 experiment, and are therefore independent from the corresponding AOGCM experiments. All values are temperature change with respect to the mean control values over the whole 150 year period.

The simple analytical model is able to reproduce the evolution of surface air temperature in response to both a step-forcing and a gradual forcing scenario. The fit seems to be very accurate to mimic the behavior of the surface temperature in a case of an abrupt forcing, not only at the beginning and at the end of the period (used in the tuning), but also in the intermediate period of transition between the two modes. However, for some models, a slight overestimation is observed for the 1% y^{-1} CO_2 scenario (CSIRO, MIROC, MPIM) and for the 2xCO_2 stabilization (GFDL). It may be due to the imperfect logarithmic dependency between the radiative forcing and the carbon dioxide concentration (e.g. because of tropospheric adjustment) or to limitations inherent to the linear two-box model such as the use of a single feedback parameter for all radiative forcing amplitude, the assumption of linearity between the radiative imbalance and the surface temperature change during a climate transition or an oversimplified representation of ocean heat uptake.

It is possible that using a median scenario to fit the EBM's parameters would give more accurate results. The abrupt 4xCO_2 case is an extreme case and an intermediate CO_2 increase scenario such as a doubling of carbon dioxide concentration may give more adequate results.

Overall, it appears that the climate response depicted by the AOGCMs can be captured by a properly-tuned two-box climate model.

c. *Upper and deep-ocean heat-uptake contributions to the fast and slow responses*

In this section, the concepts of upper and deep-ocean heat-uptake temperatures are introduced. The heat-uptake temperature T_H [Eq. 5] can be decomposed into the sum of an upper-ocean heat-uptake temperature T_U and a deep-ocean heat-uptake temperature T_D with:

$$T_U = -\frac{1}{\lambda} C \frac{dT}{dt}, \tag{23}$$

$$T_D = -\frac{1}{\lambda} C_0 \frac{dT_0}{dt}. \tag{24}$$

The contribution of these two components to the fast and the slow responses are quantitatively examined with two forcing functions.

1) Step-forcing

In the case of a step-forcing, by using Eq. (5), the heat-uptake temperature T_H is:

$$T_H(t) = -\frac{1}{\lambda} C \frac{dT}{dt} - \frac{1}{\lambda} C_0 \frac{dT_0}{dt} = -\frac{\mathcal{F}}{\lambda^2} \left[\frac{C + \phi_f C_0}{\tau_f} a_f e^{-t/\tau_f} + \frac{C + \phi_s C_0}{\tau_s} a_s e^{-t/\tau_s} \right]. \tag{25}$$

The heat-uptake temperature tends exponentially to zero with slow and fast relaxation times:

$$T_H(t) = -\frac{\mathcal{F}}{\lambda} \left[(f_U + f_D) a_f e^{-t/\tau_f} + (s_U + s_D) a_s e^{-t/\tau_s} \right], \tag{26}$$

with $f_U + f_D = 1$ and $s_U + s_D = 1$. Each mode (slow and fast) is respectively written as the sum of the contribution of each component heat uptake (the subscript U refers to the first layer and the subscript D to the second layer). Indeed, f_U (f_D) and s_U (s_D) are the partial contributions of the upper (deep) component respectively to the fast and the slow responses:

$$T_U(t) = -\frac{\mathcal{F}}{\lambda} \left[f_U a_f e^{-t/\tau_f} + s_U a_s e^{-t/\tau_s} \right], \tag{27}$$

$$T_D(t) = -\frac{\mathcal{F}}{\lambda} \left[f_D a_f e^{-t/\tau_f} + s_D a_s e^{-t/\tau_s} \right], \tag{28}$$

with:

$$f_U = \frac{C}{\lambda \tau_f} \quad ; \quad f_D = \frac{\phi_f C_0}{\lambda \tau_f} \quad ; \quad s_U = \frac{C}{\lambda \tau_s} \quad ; \quad s_D = \frac{\phi_s C_0}{\lambda \tau_s}. \tag{29}$$

2) Linear forcing

In the case of a linear forcing, the heat-uptake temperature is:

$$T_H(t) = -\frac{F}{\lambda^2}\left[C + C_0 - \sum_{i=\{s,f\}} (C + \phi_i C_0) a_i e^{-t/\tau_i}\right], \tag{30}$$

which can be rewritten as:

$$T_H(t) = -\frac{F}{\lambda}\frac{C + C_0}{\lambda}\left[h_U + h_D - (f_U + f_D) l_f e^{-t/\tau_f} - (s_U + s_D) l_s e^{-t/\tau_s}\right]. \tag{31}$$

where l_f and l_s are a fractional contribution of the fast and the slow terms: $l_f + l_s = 1$. Their expression is given in Table 1. f_U, f_D, s_U and s_D are the same as previously, $h_U = C/(C+C_0)$ and $h_D = C_0/(C + C_0)$ correspond to the fractional contribution of upper and lower layers to the asymptotic heat-uptake temperature which is proportional to the sum of the two heat capacities:

$$T_H(t) \rightarrow \hat{T}_H = -\frac{F}{\lambda}\frac{C + C_0}{\lambda}(h_U + h_D) = -\frac{F}{\lambda}\frac{C + C_0}{\lambda}. \tag{32}$$

3) Quantitative estimates of fractional contributions

Figure 5a shows the fractional contributions of the fast and slow modes to the maximum amplitude of the heat-uptake temperature for the step-forcing, a_f and a_s. For all models except one (CSIRO), the percentage of T_H due to fast response is larger than that due to the slow response for a step forcing but with a similar order of magnitude. The multimodel mean value of a_f is 59%.

The contributions of the the upper and lower layer heat uptake to the fast (f_U and f_D) and the slow (s_U and s_D) terms are depicted in Fig. 5b, c. For the fast mode, the role of the two components of the system is opposite but with a similar amplitude. For all models, the amplitude of the atmosphere/land/upper-ocean contribution T_U is larger than that of the deep ocean. For the slow mode, the contribution of T_U is negligible (i.e. $s_U \ll s_D$). Then the temperature slow response is driven exclusively by the deep-ocean heat uptake.

The fast and slow modes of the deep-ocean heat-uptake temperature T_D are of opposite sign with equal initial amplitude. During a step-forcing transient regime, T_D decreases from zero towards negative values (the heat uptake H increases from zero) until the fast mode becomes negligible. Then T_D increases slowly and tends asymptotically towards zero. This non-monotonic time evolution results from the fact that the surface and the deep-ocean temperature perturbations T and T_0 associated to the fast response have opposite signs ($\phi_f < 0$): in the fast response, the heat flux between the lower and upper layer is upward – the deep ocean warms the surface, as pointed out in Section 2.c.1.

In the case of a linear forcing, the contribution l_f of the fast term is negligible (Fig. 5d) with a multimodel mean value of 0.03%, due to the fact that the fractional amplitude l_f and l_s are proportional to their respective relaxation times. The heat-uptake temperature is driven by the deep-ocean heat-uptake temperature slow term ($s_U \ll s_D$) and by the asymptotic term $\hat{T}_H$. The upper-ocean heat-uptake temperature fast term reaches its asymptotic value which represents on average 8 % of the asymptotic heat-uptake temperature $\hat{T}_H$. Figure 5e shows the multimodel contributions to $\hat{T}_H$ of both upper ocean and deep ocean. However, on the centennial scale, the asymptotic deep-ocean heat-uptake temperature is not reached, the deep-ocean heat-uptake temperature slow term being not negligible. As a result, its contribution relatively to the upper-ocean heat-uptake temperature is smaller during the transient regime (Fig. 5f). T_U is on average 19% of the heat-uptake temperature at the time of 2xCO_2 (t=70 y) and 13% at the time of 4xCO_2 (t=140 y). Removing the upper-ocean heat-uptake contribution, the transient climate response (i.e. T at the time of 2xCO_2) would be on average 0.12% larger, which corresponds to a temperature difference on average of 0.2 K (and a range of 0.1 to 0.4 K).

4. Conclusion

In this study, we describe the analytical solutions of a two-box energy-balance model for different idealized forcings and propose a method to tune the parameters of this simple climate model to reproduce the behavior of individual coupled atmosphere-ocean general circulation models. In this simple idealized framework, the global mean surface response change consists of the sum of an instantaneous equilibrium temperature and a disequilibrium temperature, the heat-uptake temperature, which is a sum of two modes. One mode responds very quickly to changes in forcing, whereas the other mode has a larger relaxation time.

By analyzing the results of twelve AOGCMs experiments from CMIP5, we show that the calibration method based on a step-forcing scenario only allows one to derive this decomposition in two modes for any AOGCM. We first show that this decomposition can reproduce well the behavior of AOGCMs response to a step 4xCO_2 forcing scenario over the 150 year period covered by the CMIP5 simulations. We also find that the simple model calibrated from a step-forcing experiment is able to represent gradual CO_2-increase idealized scenarios because the analytic response exhibits a satisfactory fit for the 1% y^{-1} CO_2 increase scenario and stabilization when available. We found a clear separation of timescales, since the fast relaxation time multimodel mean is about 4 years while the slow timescale is about 250 years.

An analysis of the contribution of the two layers' heat uptake to the fast and the slow modes shows that the upper ocean heat uptake contributes only to the fast mode that is shown to be negligible in the case of a linear forcing. It contributes to about 20% of the deviation from equilibrium in the case of a gradual increase of the radiative perturbation. In the case of a step-forcing, both layers' heat uptake contribute to the response amplitude and the upper-ocean heat uptake plays a key role in the representation of the first stages of the temperature and radiative flux responses. Thus, this contribution is important to estimate the amplitude of the forcing from a step-forcing experiment. Moreover, an accurate representation of the temperature response near equilibrium is necessary to estimate the equilibrium climate sensitivity. The two-box EBM is the simplest tool that incorporates both of these features, and is therefore the simplest adequate model to simulate transient climate change under all kind of idealized scenarios.

However, a main limit of the simple model used in this study is the intrinsic assumption of a linear dependancy between the radiation imbalance at TOA and the mean surface temperature perturbation. In Part II, the two-box EBM with an efficacy factor of deep-ocean heat uptake proposed in Held et al. (2010) is used to overcome this problem and applied to CMIP5 AOGCMs.

Acknowledgments.

We thank Laurent Terray and Julien Boé for helpful discussions and valuable comments on the work. Thanks are also due to Isaac Held for sharing interesting ideas in his blog. This work was supported by the European Union FP7 Integrated Project COMBINE.

APPENDIX

A. General solution of the differential system

By rewriting in matrix form the set of coupled differential equations of the system [Eqs. (1) and (2)], one finds:

$$\frac{dX}{dt} = AX + B, \tag{A1}$$

with

$$X(t) = \begin{pmatrix} T \\ T_0 \end{pmatrix} \; ; \; A = \begin{bmatrix} -(\lambda+\gamma)/C & \gamma/C \\ \gamma/C_0 & -\gamma/C_0 \end{bmatrix} \; ; \; B(t) = \begin{pmatrix} \mathcal{F}/C \\ 0 \end{pmatrix}. \tag{A2}$$

The solution X^* of the homogeneous system ($B = 0$) is given by:

$$X^*(t) = e^{tA} X(0). \tag{A3}$$

Yet, A can be factorized as $A = \Phi D \Phi^{-1}$ where D is the diagonal matrix whose diagonal elements are the eigenvalues of A. One can show that:

$$D = \begin{bmatrix} -1/\tau_f & 0 \\ 0 & -1/\tau_s \end{bmatrix} \text{ and } \Phi = \begin{bmatrix} 1 & 1 \\ \phi_f & \phi_s \end{bmatrix}. \tag{A4}$$

The expression of τ_i and ϕ_i are given in Table 1. Since $e^{tA} = \Phi e^{tD} \Phi^{-1}$,

$$e^{tA} = \Phi \begin{bmatrix} e^{-t/\tau_f} & 0 \\ 0 & e^{-t/\tau_s} \end{bmatrix} \Phi^{-1}, \tag{A5}$$

and the general solution of the homogeneous system is given by:

$$T^*(t) = \frac{1}{\phi_s - \phi_f}(T_1 e^{-t/\tau_f} + T_2 e^{-t/\tau_s}), \tag{A6}$$

$$T_0^*(t) = \frac{1}{\phi_s - \phi_f}(\phi_f T_1 e^{-t/\tau_f} + \phi_s T_2 e^{-t/\tau_s}), \tag{A7}$$

with $T_1 = \phi_s T(0) - T_0(0)$ and $T_2 = -\phi_f T(0) + T_0(0)$.

To obtain the general solution of the non-homogeneous system ($B(t) \neq 0$), one can use the method known as *variation of parameter* by determining a particular solution of the form $X(t) = e^{tA}U(t)$. By noting $U'(t) = (e^{tA})^{-1}B(t)$, it is possible to derive the vector U.

Finally, for any given forcing function $t \to \mathcal{F}(t)$, the general solution of the system (A1) is given by:

$$T(t) = T^*(t) + \frac{1}{C(\phi_s - \phi_f)} \left(\phi_s \int_0^t \mathcal{F}(\xi) e^{-(t-\xi)/\tau_f} d\xi - \phi_f \int_0^t \mathcal{F}(\xi) e^{-(t-\xi)/\tau_s} d\xi \right), \tag{A8}$$

$$T_0(t) = T_0^*(t) + \frac{\phi_s \phi_f}{C(\phi_s - \phi_f)} \left(\int_0^t \mathcal{F}(\xi) e^{-(t-\xi)/\tau_f} d\xi - \int_0^t \mathcal{F}(\xi) e^{-(t-\xi)/\tau_s} d\xi \right). \tag{A9}$$

Later on, we will consider $T(0) = 0$ and $T_0(0) = 0$. So, we have $T^*(t) = T_0^*(t) = 0$.

B. Stabilization and abrupt return to preindustrial forcing

a. Linearly increasing forcing and stabilization

The GFDL provided simulations with a 1% y^{-1} CO_2 increase up to a doubling of the atmospheric CO_2 concentration followed by a stabilization of this concentration at 2xCO_2. Such a simulation was also performed with the CNRM climate model. These experiments are shown in Figs. 3 and 4. The corresponding analytical solution of the two-box model are described hereafter.

In the case of a stabilization starting from time t_{st} of a 1% y^{-1} CO_2 experiment:

$$\mathcal{F}(t) = \begin{cases} 0 & \text{if } t < 0 \\ Ft & \text{if } 0 \leq t < t_{st} \\ Ft_{st} & \text{if } t \geq t_{st}, \end{cases} \tag{A10}$$

the analytical solution for $0 \leq t < t_{st}$ is the linear-forcing solution [Eqs (14) and (15)]. For $t \geq t_{st}$, the solution is:

$$T(t) = \frac{F}{\lambda} t_{st} - \frac{F}{\lambda} \sum_{i=\{s,f\}} \tau_i a_i (1 - e^{-\frac{t_{st}}{\tau_i}}) e^{-\frac{t - t_{st}}{\tau_i}}, \tag{A11}$$

$$T_0(t) = \frac{F}{\lambda} t_{st} - \frac{F}{\lambda} \sum_{i=\{s,f\}} \phi_i \tau_i a_i (1 - e^{-\frac{t_{st}}{\tau_i}}) e^{-\frac{t - t_{st}}{\tau_i}}. \tag{A12}$$

b. *Abrupt return to preindustrial (zero) forcing*

Held et al. (2010) highlighted the interest of this case, showing that the slow response of the climate would maintain a significant climate perturbation, even if geoengineering were to provide a way to remove large amounts of CO_2 from the climate system. We hereafter describe the analytical solution corresponding to such abrupt return to preindustrial (zero) radiative forcing from a linear-forcing experiment.

In the case of an instantaneous return to preindustrial forcing at $t = t_{ar}$ from a linear-forcing transient regime:

$$\mathcal{F}(t) = \begin{cases} 0 & \text{if } t < 0 \\ Ft & \text{if } 0 \leq t < t_{ar} \\ 0 & \text{if } t \geq t_{ar}, \end{cases} \tag{A13}$$

the analytical solution for $t \geq t_{ar}$ is:

$$T(t) = \frac{F}{\lambda} \sum_{i=\{s,f\}} \tau_i a_i \left(e^{-\frac{t_{ar}}{\tau_i}} - 1 + \frac{t_{ar}}{\tau_i} \right) e^{-\frac{t - t_{ar}}{\tau_i}}, \tag{A14}$$

$$T_0(t) = \frac{F}{\lambda} \sum_{i=\{s,f\}} \phi_i \tau_i a_i \left(e^{-\frac{t_{ar}}{\tau_i}} - 1 + \frac{t_{ar}}{\tau_i} \right) e^{-\frac{t - t_{ar}}{\tau_i}}. \tag{A15}$$

When neglecting the fast term, the remaining term, which slowly tends to zero, is the *recalcitrant* component of global warming (Held et al. 2010).

C. Periodic forcing

The two-box EBM can be used to understand not only long-term climate trends due CO_2, but also to study climate perturbations due to other radiative perturbations (such as perturbations of the solar forcing), and even climate variability resulting from the variability of the radiative forcing. As an example, we hereafter give the analytical solution of the two-box EBM response to a periodic forcing, that could be used to understand the climate variability associated to the natural solar variability.

In a stationary regime, the solution of a periodic forcing $\mathcal{F}(t) = \mathcal{F}e^{i\omega t}$ is:

$$T(t) = \frac{C_0 i\omega + \gamma}{(Ci\omega + \lambda + \gamma)(C_0 i\omega + \gamma) - \gamma^2}\mathcal{F}(t), \tag{A16}$$

$$T_0(t) = \frac{\gamma}{(Ci\omega + \lambda + \gamma)(C_0 i\omega + \gamma) - \gamma^2}\mathcal{F}(t). \tag{A17}$$

The transfer function $\mathcal{H}$ of the system is thus:

$$\mathcal{H}(i\omega) = \frac{T(t)}{T_{eq}(t)} = \left[\frac{a_f}{1 + i\omega\tau_f} + \frac{a_s}{1 + i\omega\tau_s}\right]. \tag{A18}$$

We can also write the transfer function upon a canonical form:

$$\mathcal{H}(i\omega) = \frac{1 + i\frac{\omega}{\omega_1}}{\frac{(i\omega)^2}{\omega_0^2} + 2\xi\frac{i\omega}{\omega_0} + 1}, \tag{A19}$$

by noting $\omega_1 = 1/(a_f\tau_s + a_s\tau_f)$, $\omega_0 = 1/\sqrt{\tau_f\tau_s}$ and $\xi = (\tau_f + \tau_s)/(2\sqrt{\tau_f\tau_s})$.

By using the notation $\overline{\omega} = \omega/\omega_0$, the gain G of the system is given by

$$G(\omega) = |\mathcal{H}(i\omega)| = \frac{\sqrt{1 + (\omega/\omega_1)^2}}{\sqrt{(1 - \overline{\omega}^2)^2 + (2\xi\overline{\omega})^2}}. \tag{A20}$$

And the phase Φ is:

$$\Phi(\omega) = \arctan\left(\frac{\omega}{\omega_1}\right) + \arctan\left(\frac{2\xi\overline{\omega}}{\overline{\omega}^2 - 1}\right). \tag{A21}$$

The Bode diagram which plots $20\log G(\omega)$ and $\Phi(\omega)$ against $\log\omega$ is represented in Fig. 6. Asymptotically, we have:

$$G(\omega \to 0) = 0, \tag{A22}$$

$$G(\omega \to \infty) = -20\log\omega + 20\log(\omega_c), \tag{A23}$$

with a cut-off frequency

$$\omega_c = \omega_0^2/\omega_1 = \frac{\lambda}{C}. \tag{A24}$$

For the 11 year solar cycle, with $\lambda = 1.3$ W m^{-2} K^{-1}, $C = 8$ W y m^{-2} K^{-1}, $C_0 = 100$ W y m^{-2} K^{-1} and $\gamma = 0.7$ W m^{-2} K^{-1}, the amplitude of the response is attenuated by approximatively 10 dB and is shifted by about 4 years.

REFERENCES

Budyko, M. I., 1969: The effect of solar radiation variations on the climate of the Earth. *Tellus*, **5**, 611–619.

Dickinson, R., 1981: Convergence rate and stability of ocean-atmosphere coupling schemes with a zero-dimensional climate model. *J. Atmos. Sci.*, **38**, 2112–2120.

Dufresne, J.-L. and S. Bony, 2008: An assessment of the primary sources of spread of global warming estimates from coupled atmosphere-ocean models. *J. Atmos. Sci.*, **21**, 5135–5144.

Good, P., J. M. Gregory, and J. A. Lowe, 2011: A step-response simple climate model to reconstruct and interpret AOGCM projections. *Geophys. Res. Lett.*, **38**, L01 703.

Gregory, J. and M. Webb, 2008: Tropospheric adjustment induces a cloud component in CO_2 forcing. *J. Climate*, **21**, 58–71.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Clim. Dyn.*, **16**, 505–515.

Gregory, J. M. and P. M. Forster, 2008: Transient climate response estimated from radiative forcing and observed temperature change. *J. Geophys. Res.*, **113**, D23 105.

Gregory, J. M. and J. F. B. Mitchell, 1997: The climate response to CO_2 of the Hadley Centre coupled AOGCM with and without flux adjustment. *Geophys. Res. Lett.*, **24**, 1943–1946.

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.

Hansen, J., and Coauthors, 2005: Efficacy of climate forcings. *J. Geophys. Res.*, **110**, D18104, doi:10.1029/2005JD005776.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1993: On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Clim. Dyn.*, **9**, 53–61.

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate*, **23**, 2418–2427.

Intergovernmental Panel on Climate Change (IPCC), 1990: *Climate Change 1990: The Intergovernmental Panel on Climate Change Assessment*, edited by J. T. Houghton et al., Cambridge Univ. Press, New York.

Meinshausen, M., S. C. B. Raper, and T. M. L. Wigley, 2008: Emulating IPCC AR4 atmosphere-ocean and carbon cycle models for projecting global-mean, hemispheric and land/ocean temperatures: Magicc 6.0. *Atmos. Chem. Phys.*, **8**, 6153–5272.

Murphy, J. M., 1995: Transient response of the Hadley Centre coupled ocean-atmosphere model to increasing carbon dioxide. Part III: Analysis of global-mean response using simple models. *J. Climate*, **8**, 496–514.

Olivié, D. J. L., G. P. Peters, and D. Saint-Martin, 2011: On the calibration of linear response simple climate models. *submitted to J. of Climate*.

Plattner, G. K., and Coauthors, 2008: Long-Term Climate Commitments Projected with ClimateCarbon Cycle Models. *J. Climate*, **21**, 2721–2751.

Raper, S. C. B., J. M. Gregory, and R. J. Stouffer, 2002: The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *J. Climate*, **15**, 124–130.

Sellers, W. D., 1969: A global climatic model based on the energy balance of the Earth-atmosphere system. *J. Appl. Meteor.*, **8**, 392–400.

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

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2011: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, doi:10.1175/BAMS–D–11–00 094.1.

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

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

List of Tables

TABLE 1. Summary of definitions of two-box model general and mode parameters and relationships between mode and physical parameters.

Definition of general parameters	
$b = \frac{\lambda+\gamma}{C} + \frac{\gamma}{C_0}$	
$b^* = \frac{\lambda+\gamma}{C} - \frac{\gamma}{C_0}$	
$\delta = b^2 - 4\frac{\lambda\gamma}{CC_0}$	
Mode parameters	
fast	slow
$a_f = \frac{\phi_s\tau_f}{C(\phi_s-\phi_f)}\lambda$	$a_s = -\frac{\phi_f\tau_s}{C(\phi_s-\phi_f)}\lambda$
$\phi_f = \frac{C}{2\gamma}(b^* - \sqrt{\delta})$	$\phi_s = \frac{C}{2\gamma}(b^* + \sqrt{\delta})$
$\tau_f = \frac{CC_0}{2\lambda\gamma}(b - \sqrt{\delta})$	$\tau_s = \frac{CC_0}{2\lambda\gamma}(b + \sqrt{\delta})$
$l_f = a_f\tau_f\frac{\lambda}{C+C_0}$	$l_s = a_s\tau_s\frac{\lambda}{C+C_0}$
Relationships between parameters	
$a_f + a_s = 1$	
$a_f/\tau_f + a_s/\tau_s = \lambda/C$	
$\tau_f a_f + \tau_s a_s = (C + C_0)/\lambda$	
$\tau_f a_s + \tau_s a_f = C_0/\gamma$	
$\phi_f a_f/\tau_f + \phi_s a_s/\tau_s = 0$	
$\tau_f\tau_s = CC_0/(\lambda\gamma)$	
$C + \phi_f C_0 = \lambda\tau_f$	
$C + \phi_s C_0 = \lambda\tau_s$	
$\phi_f a_f + \phi_s a_s = 1$	
$\phi_f\phi_s = -C/C_0$	

TABLE 2. The $4xCO_2$ radiative forcing $\mathcal{F}_{4xCO_2}$, radiative feedback parameter λ for a CO_2 perturbation, and $4xCO_2$ equilibrium temperature T_{4xCO_2} estimates of the twelve CMIP5 models studied in this paper, and their multimodel mean and standard deviation. The version of the model used in this study is also indicated.

Model	$\mathcal{F}_{4xCO_2}$ (W m^{-2})	λ (W m^{-2} K^{-1})	T_{4xCO_2} (K)
BCC (BCC-CSM1-1)	6.7	1.21	5.6
CCCMA (CanESM2)	7.6	1.03	7.4
CNRM (CNRM-CM5.1)	7.3	1.11	6.5
CSIRO (CSIRO-Mk3-6-0)	5.1	0.61	8.3
GFDL (GFDL-ESM2M)	6.6	1.34	4.9
INM (INMCM4)	6.2	1.51	4.1
IPSL (IPSL-CM5A-LR)	6.4	0.79	8.1
MIROC (MIROC5)	8.5	1.58	5.4
MOHC (HadGEM2-ES)	5.9	0.65	9.1
MPIM (MPI-ESM-LR)	8.2	1.14	7.3
MRI (MRI-CGCM3)	6.6	1.26	5.2
NCC (NorESM1-M)	6.2	1.11	5.6
Multimodel mean	6.8	1.11	6.5
Standard deviation	1.0	0.31	1.6

TABLE 3. The atmosphere/land/upper-ocean surfacic heat capacity C, deep-ocean surfacic heat capacity C_0, heat exchange coefficient γ and fast and slow relaxation times estimates of the twelve CMIP5 models used in this paper, and their multimodel mean and standard deviation given for the 12 models ensemble and by excluding the INM model.

Model	C (W y m^{-2} K^{-1})	C_0 (W y m^{-2} K^{-1})	γ (W m^{-2} K^{-1})	τ_f (y)	τ_s (y)
BCC (BCC-CSM1-1)	7.6	53	0.67	4.0	126
CCCMA (CanESM2)	7.3	71	0.59	4.5	193
CNRM (CNRM-CM5.1)	8.4	99	0.50	5.2	289
CSIRO (CSIRO-Mk3-6-0)	6.0	69	0.88	3.9	200
GFDL (GFDL-ESM2M)	8.1	105	0.90	3.6	197
INM (INMCM4)	8.6	317	0.65	4.0	698
IPSL (IPSL-CM5A-LR)	7.7	95	0.59	5.5	286
MIROC (MIROC5)	8.3	145	0.76	3.5	285
MOHC (HadGEM2-ES)	6.5	82	0.55	5.3	280
MPIM (MPI-ESM-LR)	7.3	71	0.72	3.9	164
MRI (MRI-CGCM3)	8.5	64	0.66	4.3	150
NCC (NorESM1-M)	8.0	105	0.88	4.0	218
Multimodel mean	7.7	106	0.70	4.3	257
- without INM	7.6	87	0.70	4.3	217
Standard deviation	0.8	71	0.13	0.7	150
- without INM	0.8	26	0.14	0.7	60

List of Figures

a) One-box model b) Two-box model

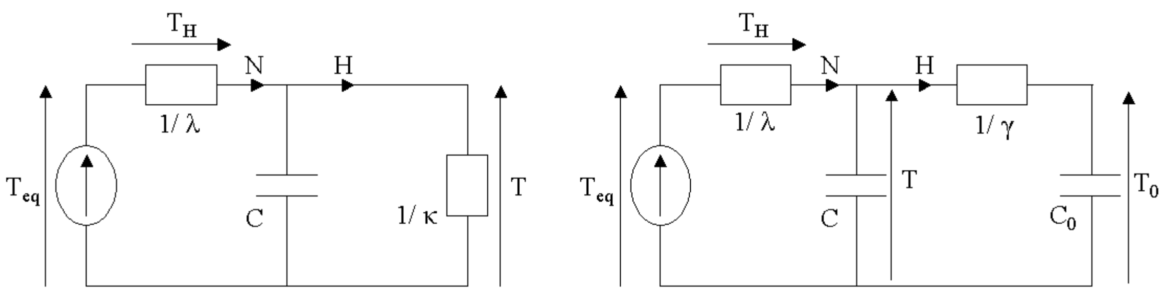

Fig. 1. Analogous electrical circuit of the one-box energy-balance model (a) and of the two-box energy-balance model (b).

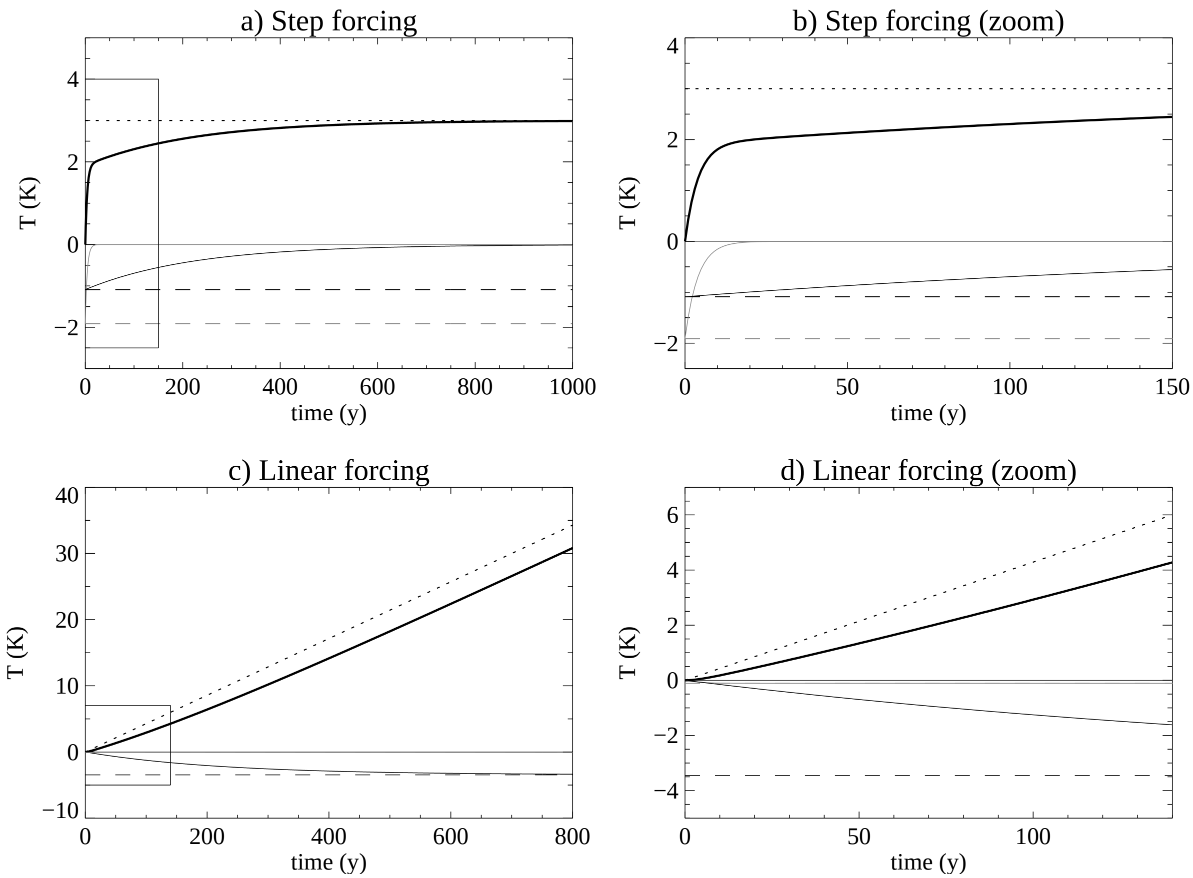

FIG. 2. Mean surface temperature response (thick black) and its decomposition in three components, the equilibrium (black dot), the fast (thin grey) and the slow (thin black) modes, as a function of time for a step-forcing (first row) and a linear forcing (second row). The dashed grey and black lines denote respectively the fast and slow modes amplitude. The right column panels are a zoom over the black box indicated on the respective left panel. Note that the fast mode and the fast mode amplitude lines are mostly merged with the y=0 line. Values are for $\mathcal{F} = 3.9$ W m^{-2}, $\lambda = 1.3$ W m^{-2} K^{-1}, $C = 8$ W y m^{-2} K^{-1}, $C_0 = 100$ W y m^{-2} K^{-1}, $\gamma = 0.7$ W m^{-2} K^{-1}.

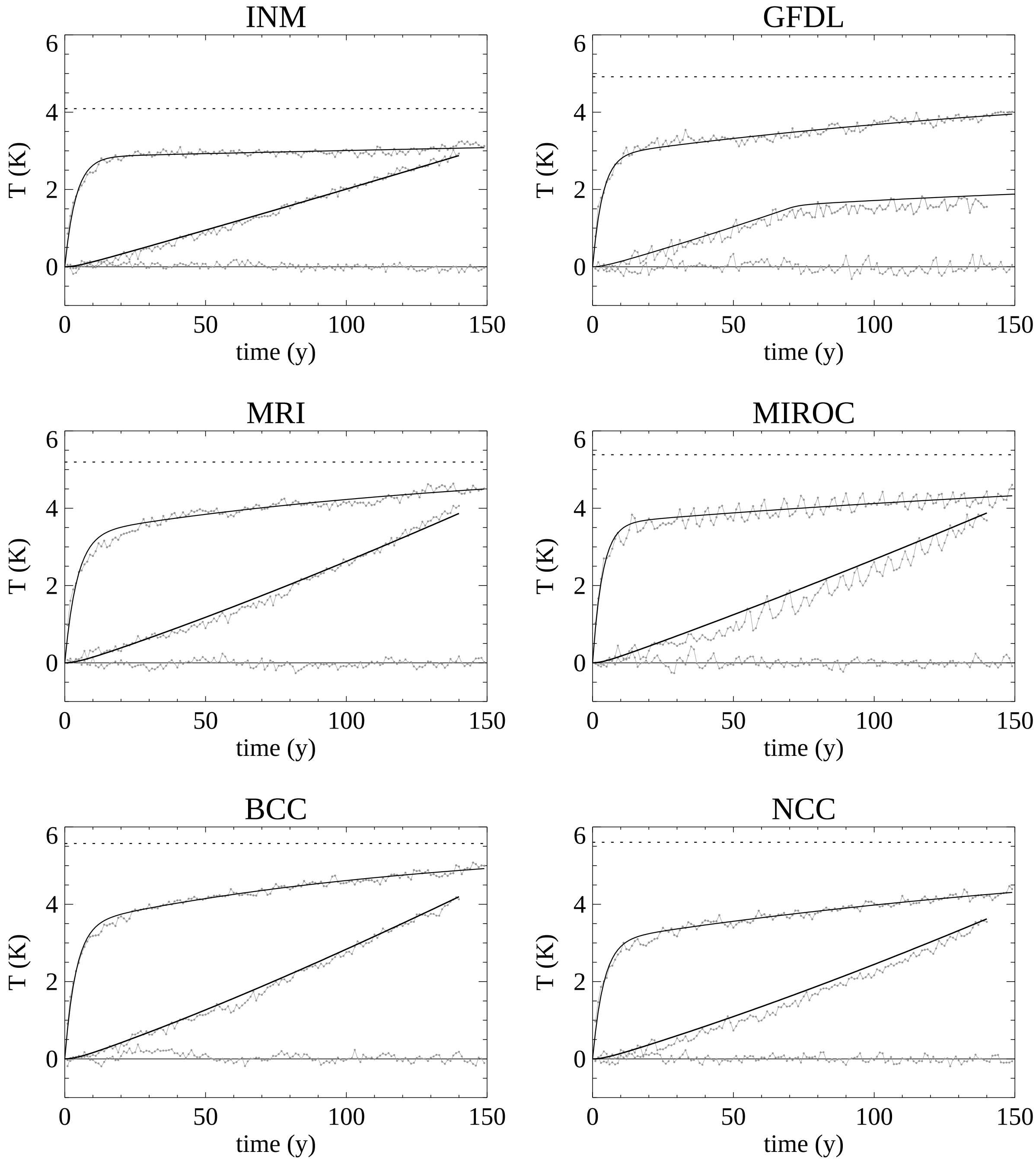

FIG. 3. Time series of global mean and annual mean surface air temperature change (gray lines) in response to the abrupt $4xCO_2$, the 1% y^{-1} CO_2 (until $4xCO_2$) and $2xCO_2$ stabilization (when available) CMIP5 experiments for the 6 AOGCMs with the lowest equilibrium climate sensitivity and of the corresponding EBM analytical temperature evolutions (black lines) calibrated from the abrupt $4xCO_2$ experiment only. For each model, the black dotted line indicates the estimated equilibrium temperature response T_{4xCO_2} for $4xCO_2$ step-forcing. All values are temperature changes with respect to the mean control value over the whole 150 years period.

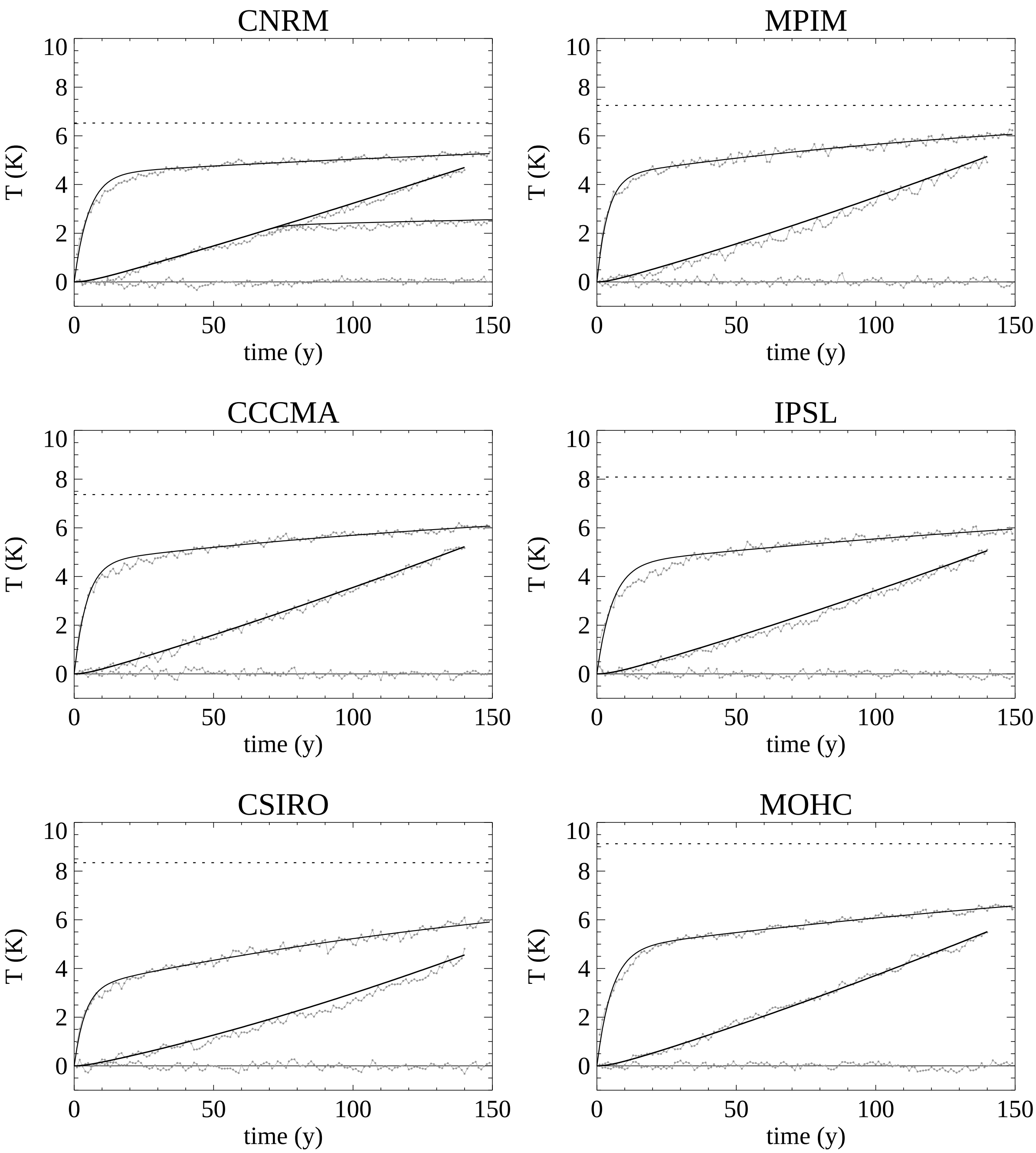

FIG. 4. Same as Fig. 3 for the 6 AOGCMs with the highest equilibrium climate sensitivity. Note that the scale of the y-axis has been modified.

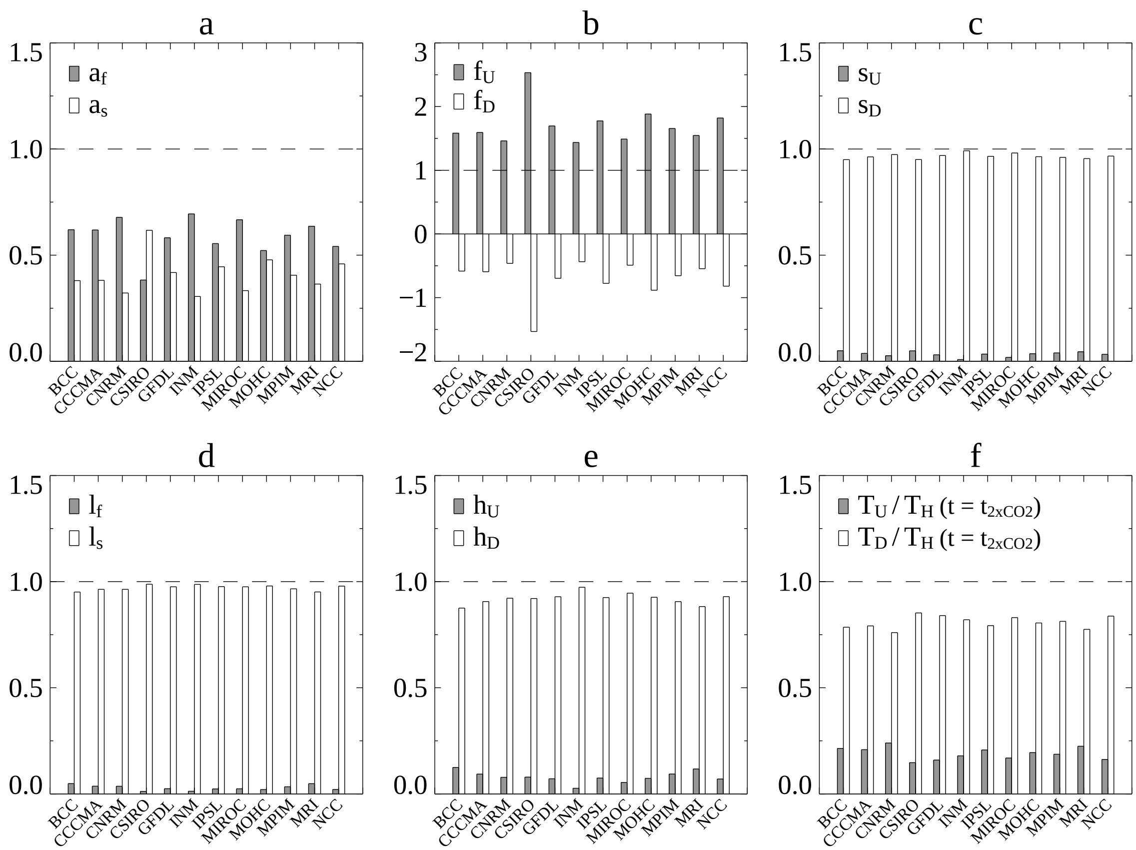

FIG. 5. Multi-model values of (a) the mode parameters (a_f, a_s), the upper-ocean and deep-ocean heat-uptake temperatures' contribution to: (b) the fast mode (f_U, f_D) and (c) the slow mode (s_U, s_D), (d) the mode parameters (l_f, l_s) in response to a linearly increasing forcing, (e) the upper-ocean and deep-ocean heat-uptake temperatures' contribution to the linear-forcing asymptotic term (h_U, h_D), (f) the ratios T_U/T_H and T_D/T_H at the time of $2xCO_2$ for the 1% y^{-1} CO_2 experiment.

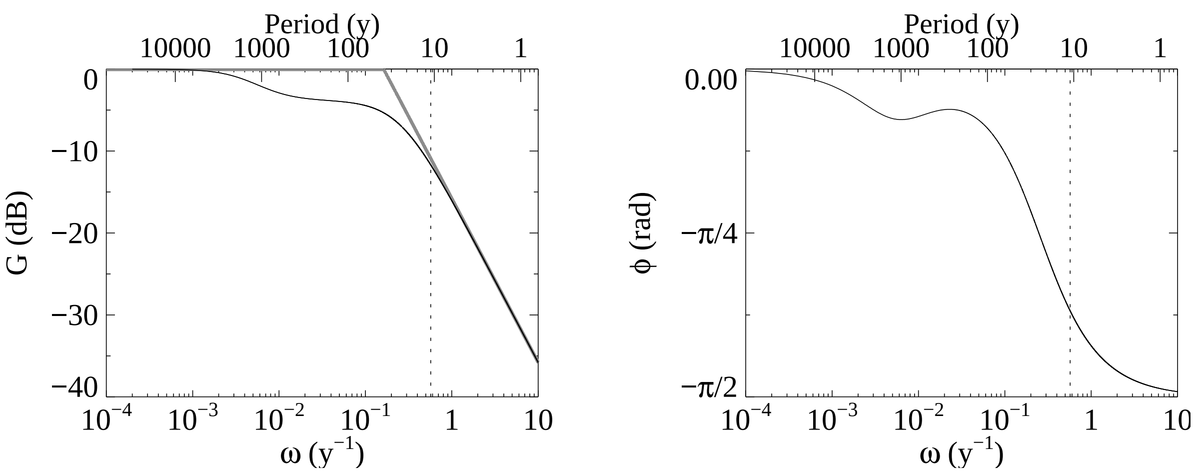

FIG. 6. Bode diagram of the the climate system in the framework of the two-box energy-balance model: gain G in decibels (left) and phase lag Φ in radians (right). For the gain plot, the values of the asymptotes (gray lines) are given in the text. The vertical dotted lines indicate a periodic forcing of 11 years. Values of gain and phase lag are for $\lambda = 1.3$ W m^{-2} K^{-1}, $C = 8$ W y m^{-2} K^{-1}, $C_0 = 100$ W y m^{-2} K^{-1}, $\gamma = 0.7$ W m^{-2} K^{-1}.